

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
25 May 2001 (25.05.2001)

PCT

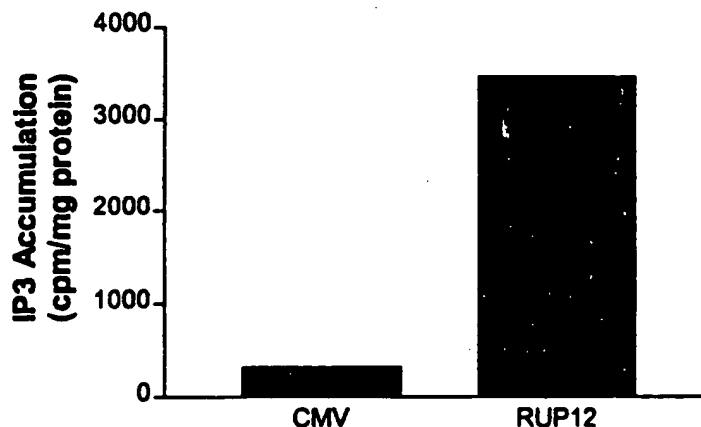
(10) International Publication Number  
WO 01/36471 A2

- (51) International Patent Classification<sup>7</sup>: C07K 14/00 60/242,332 20 October 2000 (20.10.2000) US  
60/242,343 20 October 2000 (20.10.2000) US
- (21) International Application Number: PCT/US00/31509 60/243,019 24 October 2000 (24.10.2000) US
- (22) International Filing Date: 16 November 2000 (16.11.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |            |                                |    |
|------------|--------------------------------|----|
| 60/166,088 | 17 November 1999 (17.11.1999)  | US |
| 60/166,099 | 17 November 1999 (17.11.1999)  | US |
| 60/166,369 | 17 November 1999 (17.11.1999)  | US |
| 60/171,900 | 23 December 1999 (23.12.1999)  | US |
| 60/171,901 | 23 December 1999 (23.12.1999)  | US |
| 60/171,902 | 23 December 1999 (23.12.1999)  | US |
| 60/181,749 | 11 February 2000 (11.02.2000)  | US |
| 60/189,258 | 14 March 2000 (14.03.2000)     | US |
| 60/189,259 | 14 March 2000 (14.03.2000)     | US |
| 60/195,898 | 10 April 2000 (10.04.2000)     | US |
| 60/195,899 | 10 April 2000 (10.04.2000)     | US |
| 60/196,078 | 10 April 2000 (10.04.2000)     | US |
| 60/200,419 | 28 April 2000 (28.04.2000)     | US |
| 60/203,630 | 12 May 2000 (12.05.2000)       | US |
| 60/210,741 | 12 June 2000 (12.06.2000)      | US |
| 60/210,982 | 12 June 2000 (12.06.2000)      | US |
| 60/226,760 | 21 August 2000 (21.08.2000)    | US |
| 60/235,418 | 26 September 2000 (26.09.2000) | US |
| 60/235,779 | 26 September 2000 (26.09.2000) | US |
- (71) Applicant (for all designated States except US): ARENA PHARMACEUTICALS, INC. [US/US]; 6166 Nancy Ridge Drive, San Diego, CA 92121 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): CHEN, Ruoping [CN/US]; 5296 Timber Branch Way, San Diego, CA 92130 (US). DANG, Huong, T. [US/US]; 5352 Oak Park Drive, San Diego, CA 92105 (US). LOWITZ, Kevin, P. [US/US]; 8031 Caminito de Pizza #C, San Diego, CA 82108 (US).
- (74) Agents: MILLER, Suzanne, E. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, One Liberty Place - 46th Floor, Philadelphia, PA 19103 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF HUMAN G PROTEIN-COUPLED RECEPTORS

## IP3 Assay in 293 Cells



(57) Abstract: The invention disclosed in this patent document relates to transmembrane receptors, more particularly to a human G protein-coupled receptor for which the endogenous ligand is unknown ("orphan GPCR receptors"), and most particularly to mutated (non-endogenous) versions of the human GPCRs for evidence of constitutive activity.

WO 01/36471 A2



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

- *Without international search report and to be republished upon receipt of that report.*

# ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF HUMAN G PROTEIN-COUPLED RECEPTORS

## FIELD OF THE INVENTION

5

The invention disclosed in this patent document relates to transmembrane receptors, and more particularly to human G protein-coupled receptors, and specifically to endogenous human GPCRs with particular emphasis on non-endogenous versions of the GPCRs that have been altered to establish or enhance  
10 constitutive activity of the receptor. Preferably, the altered GPCRs are used for the direct identification of candidate compounds as receptor agonists, inverse agonists or partial agonists having potential applicability as therapeutic agents.

15

## BACKGROUND OF THE INVENTION

Although a number of receptor classes exist in humans, by far the most abundant and therapeutically relevant is represented by the G protein-coupled receptor (GPCR or GPCRs) class. It is estimated that there are some 100,000 genes within the human  
20 genome, and of these, approximately 2%, or 2,000 genes, are estimated to code for GPCRs. Receptors, including GPCRs, for which the endogenous ligand has been identified are referred to as "known" receptors, while receptors for which the endogenous ligand has not been identified are referred to as "orphan" receptors. GPCRs represent an important area for the development of pharmaceutical products: from  
25 approximately 20 of the 100 known GPCRs, approximately 60% of all prescription pharmaceuticals have been developed.

GPCRs share a common structural motif. All these receptors have seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the membrane (each span is identified by number, *i.e.*,  
30 transmembrane-1 (TM-1), transmebrane-2 (TM-2), etc.). The transmembrane helices are joined by strands of amino acids between transmembrane-2 and transmembrane-3,

transmembrane-4 and transmembrane-5, and transmembrane-6 and transmembrane-7 on the exterior, or "extracellular" side, of the cell membrane (these are referred to as "extracellular" regions 1, 2 and 3 (EC-1, EC-2 and EC-3), respectively). The transmembrane helices are also joined by strands of amino acids between  
5 transmembrane-1 and transmembrane-2, transmembrane-3 and transmembrane-4, and transmembrane-5 and transmembrane-6 on the interior, or "intracellular" side, of the cell membrane (these are referred to as "intracellular" regions 1, 2 and 3 (IC-1, IC-2 and IC-3), respectively). The "carboxy" ("C") terminus of the receptor lies in the intracellular space within the cell, and the "amino" ("N") terminus of the receptor lies in the  
10 extracellular space outside of the cell.

Generally, when an endogenous ligand binds with the receptor (often referred to as "activation" of the receptor), there is a change in the conformation of the intracellular region that allows for coupling between the intracellular region and an intracellular "G-protein." It has been reported that GPCRs are "promiscuous" with respect to G proteins,  
15 *i.e.*, that a GPCR can interact with more than one G protein. *See*, Kenakin, T., 43 *Life Sciences* 1095 (1988). Although other G proteins exist, currently, Gq, Gs, Gi, Gz and Go are G proteins that have been identified. Endogenous ligand-activated GPCR coupling with the G-protein begins a signaling cascade process (referred to as "signal transduction"). Under normal conditions, signal transduction ultimately results in  
20 cellular activation or cellular inhibition. It is thought that the IC-3 loop as well as the carboxy terminus of the receptor interact with the G protein.

Under physiological conditions, GPCRs exist in the cell membrane in equilibrium between two different conformations: an "inactive" state and an "active" state. A receptor in an inactive state is unable to link to the intracellular signaling  
25 transduction pathway to produce a biological response. Changing the receptor

conformation to the active state allows linkage to the transduction pathway (via the G-protein) and produces a biological response.

A receptor may be stabilized in an active state by an endogenous ligand or a compound such as a drug. Recent discoveries, including but not exclusively limited to  
5 modifications to the amino acid sequence of the receptor, provide means other than endogenous ligands or drugs to promote and stabilize the receptor in the active state conformation. These means effectively stabilize the receptor in an active state by simulating the effect of an endogenous ligand binding to the receptor. Stabilization by such ligand-independent means is termed "constitutive receptor activation."

10

## SUMMARY OF THE INVENTION

Disclosed herein are endogenous and non-endogenous versions of human GPCRs and uses thereof.

15

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides an illustration of second messenger  $IP_3$  production from endogenous version RUP12 ("RUP12") as compared with the control ("CMV").

Figure 2 is a graphic representation of the results of a second messenger cell-based cyclic AMP assay providing comparative results for constitutive signaling of  
20 endogenous RUP13 ("RUP13") and a control vector ("CMV").

Figure 3 is a diagrammatic representation of the signal measured comparing CMV, endogenous RUP13 ("RUP13 wt") and non-endogenous, constitutively activated RUP13 ("RUP13(A268K)"), utilizing 8XCRE-Luc reporter plasmid.

Figure 4 is a graphic representation of the results of a [<sup>35</sup>S]GTPγS assay providing comparative results for constitutive signaling by RUP13:Gs Fusion Protein ("RUP13-Gs") and a control vector ("CMV").

Figure 5 is a diagrammatic representation of the signal measured comparing  
5 CMV, endogenous RUP14 ("RUP14 wt") and non-endogenous, constitutively activated RUP13 ("RUP14(L246K)"), utilizing 8XCRE-Luc reporter plasmid.

Figure 6 is a diagrammatic representation of the signal measured comparing CMV, endogenous RUP15 ("RUP15 wt") and non-endogenous, constitutively activated RUP15 ("RUP15(A398K)"), utilizing 8XCRE-Luc reporter plasmid.

10 Figure 7 is a graphic representation of the results of a second messenger cell-based cyclic AMP assay providing comparative results for constitutive signaling of endogenous RUP15 ("RUP15 wt"), non-endogenous, constitutively activated version of RUP15 ("RUP15(A398K)") and a control vector ("CMV").

Figure 8 is a graphic representation of the results of a [<sup>35</sup>S]GTPγS assay  
15 providing comparative results for constitutive signaling by RUP15:Gs Fusion Protein ("RUP15-Gs") and a control vector ("CMV").

Figure 9 provides an illustration of second messenger IP<sub>3</sub> production from endogenous version RUP17 ("RUP17") as compared with the control ("CMV").

Figure 10 provides an illustration of second messenger IP<sub>3</sub> production from  
20 endogenous version RUP21 ("RUP21") as compared with the control ("CMV").

Figure 11 is a diagrammatic representation of the signal measured comparing CMV, endogenous RUP23 ("RUP23 wt") and non-endogenous, constitutively activated RUP23 ("RUP23(W275K)"), utilizing 8XCRE-Luc reporter plasmid.

Figure 12 is a graphic representation of results from a primary screen of several candidate compounds against RUP13; results for "Compound A" are provided in well A2 and "Compound B" are provided in well G9.

5

## DETAILED DESCRIPTION

The scientific literature that has evolved around receptors has adopted a number of terms to refer to ligands having various effects on receptors. For clarity and consistency, the following definitions will be used throughout this patent document. To the extent that these definitions conflict with other definitions for these terms, the following definitions shall control:

**AGONISTS** shall mean materials (*e.g.*, ligands, candidate compounds) that activate the intracellular response when they bind to the receptor, or enhance GTP binding to membranes.

**AMINO ACID ABBREVIATIONS** used herein are set out in Table A:

**TABLE A**

ALANINE	ALA	A
ARGININE	ARG	R
ASPARAGINE	ASN	N
ASPARTIC ACID	ASP	D
CYSTEINE	CYS	C
GLUTAMIC ACID	GLU	E
GLUTAMINE	GLN	Q
GLYCINE	GLY	G
HISTIDINE	HIS	H
ISOLEUCINE	ILE	I
LEUCINE	LEU	L
LYSINE	LYS	K
METHIONINE	MET	M

PHENYLALANINE	PHE	F
PROLINE	PRO	P
SERINE	SER	S
THREONINE	THR	T
TRYPTOPHAN	TRP	W
TYROSINE	TYR	Y
VALINE	VAL	V

**PARTIAL AGONISTS** shall mean materials (e.g., ligands, candidate compounds) that activate the intracellular response when they bind to the receptor to a lesser degree/extent than do agonists, or enhance GTP binding to membranes to a lesser degree/extent than do agonists.

**ANTAGONIST** shall mean materials (e.g., ligands, candidate compounds) that competitively bind to the receptor at the same site as the agonists but which do not activate the intracellular response initiated by the active form of the receptor, and can thereby inhibit the intracellular responses by agonists or partial agonists.

10 **ANTAGONISTS** do not diminish the baseline intracellular response in the absence of an agonist or partial agonist.

**CANDIDATE COMPOUND** shall mean a molecule (for example, and not limitation, a chemical compound) that is amenable to a screening technique. Preferably, the phrase "candidate compound" does not include compounds which were publicly known to be compounds selected from the group consisting of inverse agonist, agonist or antagonist to a receptor, as previously determined by an indirect identification process ("indirectly identified compound"); more preferably, not including an indirectly identified compound which has previously been determined to have therapeutic efficacy in at least one mammal; and, most preferably, not including an indirectly identified compound which has previously been determined to have therapeutic utility in humans.

15

20



**COMPOSITION** means a material comprising at least one component; a "pharmaceutical composition" is an example of a composition.

**COMPOUND EFFICACY** shall mean a measurement of the ability of a compound to inhibit or stimulate receptor functionality, as opposed to receptor binding  
5 affinity. Exemplary means of detecting compound efficacy are disclosed in the Example section of this patent document.

**CODON** shall mean a grouping of three nucleotides (or equivalents to nucleotides) which generally comprise a nucleoside (adenosine (A), guanosine (G), cytidine (C), uridine (U) and thymidine (T)) coupled to a phosphate group and which,  
10 when translated, encodes an amino acid.

**CONSTITUTIVELY ACTIVATED RECEPTOR** shall mean a receptor subject to constitutive receptor activation. A constitutively activated receptor can be endogenous or non-endogenous.

**CONSTITUTIVE RECEPTOR ACTIVATION** shall mean stabilization of a  
15 receptor in the active state by means other than binding of the receptor with its endogenous ligand or a chemical equivalent thereof.

**CONTACT** or **CONTACTING** shall mean bringing at least two moieties together, whether in an in vitro system or an in vivo system.

**DIRECTLY IDENTIFYING** or **DIRECTLY IDENTIFIED**, in relationship to  
20 the phrase "candidate compound", shall mean the screening of a candidate compound against a constitutively activated receptor, preferably a constitutively activated orphan receptor, and most preferably against a constitutively activated G protein-coupled cell surface orphan receptor, and assessing the compound efficacy of such compound. This phrase is, under no circumstances, to be interpreted or understood to be encompassed by  
25 or to encompass the phrase "indirectly identifying" or "indirectly identified."

**ENDOGENOUS** shall mean a material that a mammal naturally produces. **ENDOGENOUS** in reference to, for example and not limitation, the term "receptor," shall mean that which is naturally produced by a mammal (for example, and not limitation, a human) or a virus. By contrast, the term **NON-ENDOGENOUS** in this context shall mean that which is not naturally produced by a mammal (for example, and not limitation, a human) or a virus. For example, and not limitation, a receptor which is not constitutively active in its endogenous form, but when manipulated becomes constitutively active, is most preferably referred to herein as a "non-endogenous, constitutively activated receptor." Both terms can be utilized to describe both "in vivo" and "in vitro" systems. For example, and not limitation, in a screening approach, the endogenous or non-endogenous receptor may be in reference to an in vitro screening system. As a further example and not limitation, where the genome of a mammal has been manipulated to include a non-endogenous constitutively activated receptor, screening of a candidate compound by means of an in vivo system is viable.

**G PROTEIN COUPLED RECEPTOR FUSION PROTEIN** and **GPCR FUSION PROTEIN**, in the context of the invention disclosed herein, each mean a non-endogenous protein comprising an endogenous, constitutively activate GPCR or a non-endogenous, constitutively activated GPCR fused to at least one G protein, most preferably the alpha ( $\alpha$ ) subunit of such G protein (this being the subunit that binds GTP), with the G protein preferably being of the same type as the G protein that naturally couples with endogenous orphan GPCR. For example, and not limitation, in an endogenous state, if the G protein "G $\alpha$ " is the predominate G protein that couples with the GPCR, a GPCR Fusion Protein based upon the specific GPCR would be a non-endogenous protein comprising the GPCR fused to G $\alpha$ ; in some circumstances, as will be set forth below, a non-predominant G protein can be fused to the GPCR. The G

protein can be fused directly to the c-terminus of the constitutively active GPCR or there may be spacers between the two.

**HOST CELL** shall mean a cell capable of having a Plasmid and/or Vector incorporated therein. In the case of a prokaryotic Host Cell, a Plasmid is typically replicated as a autonomous molecule as the Host Cell replicates (generally, the Plasmid is thereafter isolated for introduction into a eukaryotic Host Cell); in the case of a eukaryotic Host Cell, a Plasmid is integrated into the cellular DNA of the Host Cell such that when the eukaryotic Host Cell replicates, the Plasmid replicates. Preferably, for the purposes of the invention disclosed herein, the Host Cell is eukaryotic, more preferably, mammalian, and most preferably selected from the group consisting of 293, 293T and COS-7 cells.

**INDIRECTLY IDENTIFYING** or **INDIRECTLY IDENTIFIED** means the traditional approach to the drug discovery process involving identification of an endogenous ligand specific for an endogenous receptor, screening of candidate compounds against the receptor for determination of those which interfere and/or compete with the ligand-receptor interaction, and assessing the efficacy of the compound for affecting at least one second messenger pathway associated with the activated receptor.

**INHIBIT** or **INHIBITING**, in relationship to the term "response" shall mean that a response is decreased or prevented in the presence of a compound as opposed to in the absence of the compound.

**INVERSE AGONISTS** shall mean materials (*e.g.*, ligand, candidate compound) which bind to either the endogenous form of the receptor or to the constitutively activated form of the receptor, and which inhibit the baseline intracellular response initiated by the active form of the receptor below the normal base level of activity which

is observed in the absence of agonists or partial agonists, or decrease GTP binding to membranes. Preferably, the baseline intracellular response is inhibited in the presence of the inverse agonist by at least 30%, more preferably by at least 50%, and most preferably by at least 75%, as compared with the baseline response in the absence of the inverse  
5 agonist.

**KNOWN RECEPTOR** shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has been identified.

**LIGAND** shall mean an endogenous, naturally occurring molecule specific for an endogenous, naturally occurring receptor.

10 **MUTANT** or **MUTATION** in reference to an endogenous receptor's nucleic acid and/or amino acid sequence shall mean a specified change or changes to such endogenous sequences such that a mutated form of an endogenous, non-constitutively activated receptor evidences constitutive activation of the receptor. In terms of equivalents to specific sequences, a subsequent mutated form of a human receptor is  
15 considered to be equivalent to a first mutation of the human receptor if (a) the level of constitutive activation of the subsequent mutated form of a human receptor is substantially the same as that evidenced by the first mutation of the receptor; and (b) the percent sequence (amino acid and/or nucleic acid) homology between the subsequent mutated form of the receptor and the first mutation of the receptor is at least about 80%,  
20 more preferably at least about 90% and most preferably at least 95%. Ideally, and owing to the fact that the most preferred cassettes disclosed herein for achieving constitutive activation includes a single amino acid and/or codon change between the endogenous and the non-endogenous forms of the GPCR, the percent sequence homology should be at least 98%.

**NON-ORPHAN RECEPTOR** shall mean an endogenous naturally occurring molecule specific for an endogenous naturally occurring ligand wherein the binding of a ligand to a receptor activates an intracellular signaling pathway.

**ORPHAN RECEPTOR** shall mean an endogenous receptor for which the  
5 endogenous ligand specific for that receptor has not been identified or is not known.

**PHARMACEUTICAL COMPOSITION** shall mean a composition comprising at least one active ingredient, whereby the composition is amenable to investigation for a specified, efficacious outcome in a mammal (for example, and not limitation, a human). Those of ordinary skill in the art will understand and appreciate the  
10 techniques appropriate for determining whether an active ingredient has a desired efficacious outcome based upon the needs of the artisan.

**PLASMID** shall mean the combination of a Vector and cDNA. Generally, a Plasmid is introduced into a Host Cell for the purposes of replication and/or expression of the cDNA as a protein.

**SECOND MESSENGER** shall mean an intracellular response produced as a  
15 result of receptor activation. A second messenger can include, for example, inositol triphosphate (IP<sub>3</sub>), diacylglycerol (DAG), cyclic AMP (cAMP), and cyclic GMP (cGMP). Second messenger response can be measured for a determination of receptor activation. In addition, second messenger response can be measured for the direct  
20 identification of candidate compounds, including for example, inverse agonists, agonists, partial agonists and antagonists.

**STIMULATE** or **STIMULATING**, in relationship to the term "response" shall mean that a response is increased in the presence of a compound as opposed to in the absence of the compound.

**VECTOR** in reference to cDNA shall mean a circular DNA capable of incorporating at least one cDNA and capable of incorporation into a Host Cell.

The order of the following sections is set forth for presentational efficiency and is not intended, nor should be construed, as a limitation on the disclosure or the claims to follow.

#### **A. Introduction**

The traditional study of receptors has always proceeded from the a priori assumption (historically based) that the endogenous ligand must first be identified before discovery could proceed to find antagonists and other molecules that could affect the receptor. Even in cases where an antagonist might have been known first, the search immediately extended to looking for the endogenous ligand. This mode of thinking has persisted in receptor research even after the discovery of constitutively activated receptors. What has not been heretofore recognized is that it is the active state of the receptor that is most useful for discovering agonists, partial agonists, and inverse agonists of the receptor. For those diseases which result from an overly active receptor or an under-active receptor, what is desired in a therapeutic drug is a compound which acts to diminish the active state of a receptor or enhance the activity of the receptor, respectively, not necessarily a drug which is an antagonist to the endogenous ligand. This is because a compound that reduces or enhances the activity of the active receptor state need not bind at the same site as the endogenous ligand. Thus, as taught by a method of this invention, any search for therapeutic compounds should start by screening compounds against the ligand-independent active state.

#### **B. Identification of Human GPCRs**

The efforts of the Human Genome project has led to the identification of a plethora of information regarding nucleic acid sequences located within the human genome; it has been the case in this endeavor that genetic sequence information has been made available without an understanding or recognition as to whether or not any particular genomic sequence does or may contain open-reading frame information that translate human proteins. Several methods of identifying nucleic acid sequences within the human genome are within the purview of those having ordinary skill in the art. For example, and not limitation, a variety of human GPCRs, disclosed herein, were discovered by reviewing the GenBank™ database. Table B, below, lists several endogenous GPCRs that we have discovered, along with other GPCR's that are homologous to the disclosed GPCR.

**TABLE B**

<b>Disclosed Human Orphan GPCRs</b>	<b>Accession Number Identified</b>	<b>Open Reading Frame (Base Pairs)</b>	<b>Reference To Homologous GPCR</b>	<b>Per Cent Homology To Designated GPCR</b>
<b>hRUP8</b>	AL121755	1,152bp	NPY2R	27%
<b>hRUP9</b>	AC0113375	1,260bp	GAL2R	22%
<b>hRUP10</b>	AC008745	1,014bp	C5aR	40%
<b>hRUP11</b>	AC013396	1,272bp	HM74	36%
<b>hRUP12</b>	AP000808	966bp	Mas1	34%
<b>hRUP13</b>	AC011780	1,356bp	Fish GPRX-ORYLA	43%
<b>hRUP14</b>	AL137118	1,041bp	CysLT1R	35%
<b>hRUP15</b>	AL016468	1,527bp	RE2	30%
<b>hRUP16</b>	AL136106	1,068bp	GLR101	37%
<b>hRUP17</b>	AC023078	969bp	Mas1	37%
<b>hRUP18</b>	AC008547	1,305bp	Oxytocin	31%
<b>hRUP19</b>	AC026331	1,041bp	HM74	52%
<b>hRUP20</b>	AL161458	1,011bp	GPR34	25%
<b>hRUP21</b>	AC026756	1,014bp	P2Y1R	37%
<b>hRUP22</b>	AC027026	993bp	RUP17 Mas1	67% 37%

<b>hRUP23</b>	AC007104	1,092bp	Rat GPR26	31%
<b>hRUP24</b>	AL355388	1,125bp	SALPR	44%
<b>hRUP25</b>	AC026331	1,092bp	HM74	95%
<b>hRUP26</b>	AC023040	1,044bp	Rabbit 5HT1D	27%
<b>hRUP27</b>	AC027643	158,700	MCH	38%

Receptor homology is useful in terms of gaining an appreciation of a role of the receptors within the human body. As the patent document progresses, we will disclose techniques for mutating these receptors to establish non-endogenous, constitutively  
5 activated versions of these receptors.

The techniques disclosed herein have also been applied to other human, orphan GPCRs known to the art, as will be apparent as the patent document progresses.

### **C. Receptor Screening**

10 Screening candidate compounds against a non-endogenous, constitutively activated version of the human GPCRs disclosed herein allows for the direct identification of candidate compounds which act at this cell surface receptor, without requiring use of the receptor's endogenous ligand. Using routine, and often commercially available techniques, one can determine areas within the body where the  
15 endogenous version of human GPCRs disclosed herein is expressed and/or over-expressed. It is also possible using these techniques to determine related disease/disorder states which are associated with the expression and/or over-expression of the receptor; such an approach is disclosed in this patent document.

With respect to creation of a mutation that may evidence constitutive activation  
20 of the human GPCR disclosed herein is based upon the distance from the proline residue at which is presumed to be located within TM6 of the GPCR; this algorithmic technique is disclosed in co-pending and commonly assigned patent document PCT Application



Number PCT/US99/23938, published as WO 00/22129 on April 20, 2000, which, along with the other patent documents listed herein, is incorporated herein by reference. The algorithmic technique is not predicated upon traditional sequence "alignment" but rather a specified distance from the aforementioned TM6 proline residue (or, of course, endogenous constitutive substitution for such proline residue). By mutating the amino acid residue located 16 amino acid residues from this residue (presumably located in the IC3 region of the receptor) to, most preferably, a lysine residue, such activation may be obtained. Other amino acid residues may be useful in the mutation at this position to achieve this objective.

10

#### **D. Disease/Disorder Identification and/or Selection**

As will be set forth in greater detail below, most preferably inverse agonists and agonists to the non-endogenous, constitutively activated GPCR can be identified by the methodologies of this invention. Such inverse agonists and agonists are ideal candidates as lead compounds in drug discovery programs for treating diseases related to this receptor. Because of the ability to directly identify inverse agonists to the GPCR, thereby allowing for the development of pharmaceutical compositions, a search for diseases and disorders associated with the GPCR is relevant. For example, scanning both diseased and normal tissue samples for the presence of the GPCR now becomes more than an academic exercise or one which might be pursued along the path of identifying an endogenous ligand to the specific GPCR. Tissue scans can be conducted across a broad range of healthy and diseased tissues. Such tissue scans provide a preferred first step in associating a specific receptor with a disease and/or disorder.

Preferably, the DNA sequence of the human GPCR is used to make a probe for (a) dot-blot analysis against tissue-mRNA, and/or (b) RT-PCR identification of the expression of the receptor in tissue samples. The presence of a receptor in a tissue

source, or a diseased tissue, or the presence of the receptor at elevated concentrations in diseased tissue compared to a normal tissue, can be preferably utilized to identify a correlation with a treatment regimen, including but not limited to, a disease associated with that disease. Receptors can equally well be localized to regions of organs by this technique. Based on the known functions of the specific tissues to which the receptor is localized, the putative functional role of the receptor can be deduced.

## **E. Screening of Candidate Compounds**

### **1. Generic GPCR screening assay techniques**

When a G protein receptor becomes constitutively active, it binds to a G protein (e.g., Gq, Gs, Gi, Gz, Go) and stimulates the binding of GTP to the G protein. The G protein then acts as a GTPase and slowly hydrolyzes the GTP to GDP, whereby the receptor, under normal conditions, becomes deactivated. However, constitutively activated receptors continue to exchange GDP to GTP. A non-hydrolyzable analog of GTP, [<sup>35</sup>S]GTPγS, can be used to monitor enhanced binding to membranes which express constitutively activated receptors. It is reported that [<sup>35</sup>S]GTPγS can be used to monitor G protein coupling to membranes in the absence and presence of ligand. An example of this monitoring, among other examples well-known and available to those in the art, was reported by Traynor and Nahorski in 1995. The preferred use of this assay system is for initial screening of candidate compounds because the system is generically applicable to all G protein-coupled receptors regardless of the particular G protein that interacts with the intracellular domain of the receptor.

### **2. Specific GPCR screening assay techniques**

Once candidate compounds are identified using the "generic" G protein-coupled receptor assay (i.e., an assay to select compounds that are agonists, partial agonists, or inverse agonists), further screening to confirm that the compounds have interacted at the

receptor site is preferred. For example, a compound identified by the "generic" assay may not bind to the receptor, but may instead merely "uncouple" the G protein from the intracellular domain.

*a. Gs, Gz and Gi.*

- 5 Gs stimulates the enzyme adenylyl cyclase. Gi (and Gz and Go), on the other hand, inhibit this enzyme. Adenylyl cyclase catalyzes the conversion of ATP to cAMP; thus, constitutively activated GPCRs that couple the Gs protein are associated with increased cellular levels of cAMP. On the other hand, constitutively activated GPCRs that couple Gi (or Gz, Go) protein are associated with decreased cellular levels of cAMP.
- 10 *See, generally, "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3<sup>rd</sup> Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992).* Thus, assays that detect cAMP can be utilized to determine if a candidate compound is, *e.g.*, an inverse agonist to the receptor (*i.e.*, such a compound would decrease the levels of cAMP). A variety of approaches known in the art for measuring cAMP can be utilized;
- 15 a most preferred approach relies upon the use of anti-cAMP antibodies in an ELISA-based format. Another type of assay that can be utilized is a whole cell second messenger reporter system assay. Promoters on genes drive the expression of the proteins that a particular gene encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-responsive DNA binding protein or transcription
- 20 factor (CREB) that then binds to the promoter at specific sites called cAMP response elements and drives the expression of the gene. Reporter systems can be constructed which have a promoter containing multiple cAMP response elements before the reporter gene, *e.g.*,  $\beta$ -galactosidase or luciferase. Thus, a constitutively activated Gs-linked receptor causes the accumulation of cAMP that then activates the gene and expression of

the reporter protein. The reporter protein such as  $\beta$ -galactosidase or luciferase can then be detected using standard biochemical assays (Chen et al. 1995).

**b. Go and Gq.**

5 Gq and Go are associated with activation of the enzyme phospholipase C, which in turn hydrolyzes the phospholipid  $PIP_2$ , releasing two intracellular messengers: diacylglycerol (DAG) and inistol 1,4,5-triphoisphate ( $IP_3$ ). Increased accumulation of  $IP_3$  is associated with activation of Gq- and Go-associated receptors. *See, generally,* "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3<sup>rd</sup> Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). Assays that detect  $IP_3$  accumulation can be utilized to determine if a candidate compound is, *e.g.*, an inverse agonist to a Gq- or Go-associated receptor (*i.e.*, such a compound would decrease the levels of  $IP_3$ ). Gq-associated receptors can also been examined using an AP1 reporter assay in that Gq-dependent phospholipase C causes activation of genes containing AP1  
10 elements; thus, activated Gq-associated receptors will evidence an increase in the expression of such genes, whereby inverse agonists thereto will evidence a decrease in such expression, and agonists will evidence an increase in such expression. Commercially available assays for such detection are available.

**3. GPCR Fusion Protein**

20 The use of an endogenous, constitutively activate orphan GPCR or a non-endogenous, constitutively activated orphan GPCR, for use in screening of candidate compounds for the direct identification of inverse agonists, agonists and partial agonists provide an interesting screening challenge in that, by definition, the receptor is active even in the absence of an endogenous ligand bound thereto. Thus, in order to  
25 differentiate between, *e.g.*, the non-endogenous receptor in the presence of a candidate compound and the non-endogenous receptor in the absence of that compound, with an

aim of such a differentiation to allow for an understanding as to whether such compound may be an inverse agonist, agonist, partial agonist or have no affect on such a receptor, it is preferred that an approach be utilized that can enhance such differentiation. A preferred approach is the use of a GPCR Fusion Protein.

5           Generally, once it is determined that a non-endogenous orphan GPCR has been constitutively activated using the assay techniques set forth above (as well as others), it is possible to determine the predominant G protein that couples with the endogenous GPCR. Coupling of the G protein to the GPCR provides a signaling pathway that can be assessed. Because it is most preferred that screening take place by use of a mammalian  
10       expression system, such a system will be expected to have endogenous G protein therein. Thus, by definition, in such a system, the non-endogenous, constitutively activated orphan GPCR will continuously signal. In this regard, it is preferred that this signal be enhanced such that in the presence of, e.g., an inverse agonist to the receptor, it is more likely that it will be able to more readily differentiate, particularly in the context of  
15       screening, between the receptor when it is contacted with the inverse agonist.

          The GPCR Fusion Protein is intended to enhance the efficacy of G protein coupling with the non-endogenous GPCR. The GPCR Fusion Protein is preferred for screening with a non-endogenous, constitutively activated GPCR because such an approach increases the signal that is most preferably utilized in such screening  
20       techniques. This is important in facilitating a significant "signal to noise" ratio; such a significant ratio is import preferred for the screening of candidate compounds as disclosed herein.

          The construction of a construct useful for expression of a GPCR Fusion Protein is within the purview of those having ordinary skill in the art. Commercially available  
25       expression vectors and systems offer a variety of approaches that can fit the particular

needs of an investigator. The criteria of importance for such a GPCR Fusion Protein construct is that the endogenous GPCR sequence and the G protein sequence both be in-frame (preferably, the sequence for the endogenous GPCR is upstream of the G protein sequence) and that the "stop" codon of the GPCR must be deleted or replaced such that upon expression of the GPCR, the G protein can also be expressed. The GPCR can be linked directly to the G protein, or there can be spacer residues between the two (preferably, no more than about 12, although this number can be readily ascertained by one of ordinary skill in the art). We have a preference (based upon convenience) of use of a spacer in that some restriction sites that are not used will, effectively, upon expression, become a spacer. Most preferably, the G protein that couples to the non-endogenous GPCR will have been identified prior to the creation of the GPCR Fusion Protein construct. Because there are only a few G proteins that have been identified, it is preferred that a construct comprising the sequence of the G protein (*i.e.*, a universal G protein construct) be available for insertion of an endogenous GPCR sequence therein; this provides for efficiency in the context of large-scale screening of a variety of different endogenous GPCRs having different sequences.

As noted above, constitutively activated GPCRs that couple to  $G_i$ ,  $G_z$  and  $G_o$  are expected to inhibit the formation of cAMP making assays based upon these types of GPCRs challenging (*i.e.*, the cAMP signal decreases upon activation thus making the direct identification of, *e.g.*, inverse agonists (which would further decrease this signal), interesting. As will be disclosed herein, we have ascertained that for these types of receptors, it is possible to create a GPCR Fusion Protein that is not based upon the endogenous GPCR's endogenous G protein, in an effort to establish a viable cyclase-based assay. Thus, for example, an endogenous  $G_i$  coupled receptor can be fused to a  $G_s$  protein – we believe that such a fusion construct, upon expression, "drives" or "forces"

the endogenous GPCR to couple with, *e.g.*, Gs rather than the "natural" Gi protein, such that a cyclase-based assay can be established. Thus, for Gi, Gz and Go coupled receptors, we prefer that that when a GPCR Fusion Protein is used and the assay is based upon detection of adenylyl cyclase activity, that the fusion construct be established with

5 Gs (or an equivalent G protein that stimulates the formation of the enzyme adenylyl cyclase).

Equally effective is a G Protein Fusion construct that utilizes a Gq Protein fused with a Gs, Gi, Gz or Go Protein. A most preferred fusion construct can be accomplished with a Gq Protein wherein the first six (6) amino acids of the G-protein  $\alpha$ -subunit

10 ("G $\alpha$ q") is deleted and the last five (5) amino acids at the C-terminal end of G $\alpha$ q is replaced with the corresponding amino acids of the G $\alpha$  of the G protein of interest. For example, a fusion construct can have a Gq (6 amino acid deletion) fused with a Gi Protein, resulting in a "Gq/Gi Fusion Construct". We believe that this fusion construct will force the endogenous Gi coupled receptor to couple to its non-endogenous G

15 protein, Gq, such that the second messenger, for example, inositol triphosphate or diacylglycerol, can be measured in lieu of cAMP production.

#### **4. Co-transfection of a Target Gi Coupled GPCR with a Signal-Enhancer Gs Coupled GPCR (cAMP Based Assays)**

20 A Gi coupled receptor is known to inhibit adenylyl cyclase, and, therefore, decrease the level of cAMP production, which can make assessment of cAMP levels challenging. An effective technique in measuring the decrease in production of cAMP as an indication of constitutive activation of a receptor that predominantly couples Gi upon

25 activation can be accomplished by co-transfecting a signal enhancer, *e.g.*, a non-endogenous, constitutively activated receptor that predominantly couples with Gs upon activation (*e.g.*, TSHR-A623I, disclosed below), with the Gi linked GPCR. As is

apparent, constitutive activation of a Gs coupled receptor can be determined based upon an increase in production of cAMP. Constitutive activation of a Gi coupled receptor leads to a decrease in production cAMP. Thus, the co-transfection approach is intended to advantageously exploit these "opposite" affects. For example, co-transfection of a non-endogenous, constitutively activated Gs coupled receptor (the "signal enhancer") with the endogenous Gi coupled receptor (the "target receptor") provides a baseline cAMP signal (*i.e.*, although the Gi coupled receptor will decrease cAMP levels, this "decrease" will be relative to the substantial increase in cAMP levels established by constitutively activated Gs coupled signal enhancer). By then co-transfecting the signal enhancer with a constitutively activated version of the target receptor, cAMP would be expected to further decrease (relative to base line) due to the increased functional activity of the Gi target (*i.e.*, which decreases cAMP).

Screening of candidate compounds using a cAMP based assay can then be accomplished, with two provisos: first, relative to the Gi coupled target receptor, "opposite" effects will result, *i.e.*, an inverse agonist of the Gi coupled target receptor will increase the measured cAMP signal, while an agonist of the Gi coupled target receptor will decrease this signal; second, as would be apparent, candidate compounds that are directly identified using this approach should be assessed independently to ensure that these do not target the signal enhancing receptor (this can be done prior to or after screening against the co-transfected receptors).

#### **F. Medicinal Chemistry**

Generally, but not always, direct identification of candidate compounds is preferably conducted in conjunction with compounds generated via combinatorial chemistry techniques, whereby thousands of compounds are randomly prepared for such analysis. Generally, the results of such screening will be compounds having



unique core structures; thereafter, these compounds are preferably subjected to additional chemical modification around a preferred core structure(s) to further enhance the medicinal properties thereof. Such techniques are known to those in the art and will not be addressed in detail in this patent document.

5

#### **G. Pharmaceutical compositions**

Candidate compounds selected for further development can be formulated into pharmaceutical compositions using techniques well known to those in the art. Suitable pharmaceutically-acceptable carriers are available to those in the art; for example, see  
10 Remington's Pharmaceutical Sciences, 16<sup>th</sup> Edition, 1980, Mack Publishing Co., (Oslo et al., eds.).

#### **H. Other Utility**

Although a preferred use of the non-endogenous versions the human GPCRs  
15 disclosed herein may be for the direct identification of candidate compounds as inverse agonists, agonists or partial agonists (preferably for use as pharmaceutical agents), these versions of human GPCRs can also be utilized in research settings. For example, *in vitro* and *in vivo* systems incorporating GPCRs can be utilized to further elucidate and understand the roles these receptors play in the human condition, both normal and  
20 diseased, as well as understanding the role of constitutive activation as it applies to understanding the signaling cascade. The value in non-endogenous human GPCRs is that their utility as a research tool is enhanced in that, because of their unique features, non-endogenous human GPCRs can be used to understand the role of these receptors in the human body before the endogenous ligand therefore is identified. Other uses of the  
25 disclosed receptors will become apparent to those in the art based upon, *inter alia*, a review of this patent document.

## EXAMPLES

The following examples are presented for purposes of elucidation, and not limitation, of the present invention. While specific nucleic acid and amino acid sequences are disclosed herein, those of ordinary skill in the art are credited with the ability to make minor modifications to these sequences while achieving the same or substantially similar results reported below. The traditional approach to application or understanding of sequence cassettes from one sequence to another (*e.g.* from rat receptor to human receptor or from human receptor A to human receptor B) is generally predicated upon sequence alignment techniques whereby the sequences are aligned in an effort to determine areas of commonality. The mutational approach disclosed herein does not rely upon this approach but is instead based upon an algorithmic approach and a positional distance from a conserved proline residue located within the TM6 region of human GPCRs. Once this approach is secured, those in the art are credited with the ability to make minor modifications thereto to achieve substantially the same results (*i.e.*, constitutive activation) disclosed herein. Such modified approaches are considered within the purview of this disclosure.

//

//

//

20 //

### **Example 1**

#### **ENDOGENOUS HUMAN GPCRS**

##### **1. Identification of Human GPCRs**

The disclosed endogenous human GPCRs were identified based upon a review of the GenBank™ database information. While searching the database, the following cDNA clones were identified as evidenced below (Table C).

TABLE C

Disclosed Human Orphan GPCRs	Accession Number Identified	Complete DNA Sequence (Base Pairs)	Open Reading Frame (Base Pairs)	Nucleic Acid SEQ.ID. NO.	Amino Acid SEQ.ID. NO.
hRUP8	AL121755	147,566bp	1,152bp	1	2
hRUP9	AC0113375	143,181bp	1,260bp	3	4
hRUP10	AC008745	94,194bp	1,014bp	5	6
hRUP11	AC013396	155,086bp	1,272bp	7	8
hRUP12	AP000808	177,764bp	966bp	9	10
hRUP13	AC011780	167,819bp	1,356bp	11	12
hRUP14	AL137118	168,297bp	1,041bp	13	14
hRUP15	AL016468	138,828bp	1,527bp	15	16
hRUP16	AL136106	208,042bp	1,068bp	17	18
hRUP17	AC023078	161,735bp	969bp	19	20
hRUP18	AC008547	117,304bp	1,305bp	21	22
hRUP19	AC026331	145,183bp	1,041bp	23	24
hRUP20	AL161458	163,511bp	1,011bp	25	26
hRUP21	AC026756	156,534bp	1,014bp	27	28
hRUP22	AC027026	151,811bp	993bp	29	30
hRUP23	AC007104	200,000bp	1,092bp	31	32
hRUP24	AL355388	190,538bp	1,125bp	33	34
hRUP25	AC026331	145,183bp	1,092bp	35	36
hRUP26	AC023040	178,508bp	1,044bp	37	38
hRUP27	AC027643	158,700bp	1,020bp	39	40

5

## 2. Full Length Cloning

### a. hRUP8 (Seq. Id. Nos. 1 & 2)

The disclosed human RUP8 was identified based upon the use of EST database (*dbEST*) information. While searching the *dbEST*, a cDNA clone with accession number

AL121755 was identified to encode a novel GPCR. The following PCR primers were used for RT-PCR with human testis Marathon-Ready cDNA (Clontech) as templates:

5'-CTTGCAGACATCACCATGGCAGCC-3' (SEQ.ID.NO.:41; sense) and

5'-GTGATGCTCTGAGTACTGGACTGG-3' (SEQ.ID.NO.: 42; antisense).

- 5 PCR was performed using Advantage cDNA polymerase (Clontech; manufacturing instructions will be followed) in 50ul reaction by the following cycles: 94°C for 30 sec; 94°C for 10 sec; 65°C for 20 sec, 72°C for 1.5 min, and 72°C for 7 min. Cycles 2 through 4 were repeated 35 times.

A 1.2kb PCR fragment was isolated and cloned into the pCRII-TOPO vector  
10 (Invitrogen) and sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem).  
See, SEQ.ID.NO.:1. The putative amino acid sequence for RUP8 is set forth in  
SEQ.ID.NO.:2.

**b. hRUP9 (Seq. Id. Nos. 3 & 4)**

The disclosed human RUP9 was identified based upon the use of GeneBank  
15 database information. While searching the database, a cDNA clone with Accession  
Number AC011375 was identified as a human genomic sequence from chromosome  
5. The full length RUP9 was cloned by PCR using primers:  
5'-GAAGCTGTGAAGAGTGATGC-3' (SEQ.ID.NO.:43; sense),  
5'-GTCAGCAATATTGATAAGCAGCAG-3' (SEQ.ID.NO.:44; antisense)  
20 and human genomic DNA (Promega) as a template. Taq Plus Precision polymerase  
(Stratagene) was used for the amplification in a 100µl reaction with 5% DMSO by the  
following cycle with step 2 to step 4 repeated 35 times: 94°C for 1 minute; 94°C for  
30 seconds; 56°C for 30 seconds; 72°C for 2 minutes; 72°C for 5 minutes.

A 1.3 Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector  
25 (Invitrogen) from 1% agarose gel and completely sequenced using the ABI Big Dye

Terminator kit (P.E. Biosystem). See, SEQ.ID.NO.:3. The putative amino acid sequence for RUP8 is set forth in SEQ.ID.NO.:4. The sequence of RUP9 clones isolated from human genomic DNA matched with the sequence obtained from data base.

**c. hRUP10 (Seq. Id. Nos. 5 & 6)**

5           The disclosed human RUP10 was identified based upon the use of GenBank database information. While searching the database, a cDNA clone with accession number AC008754 was identified as a human genomic sequence from chromosome 19. The full length RUP10 was cloned by RT-PCR using primers:  
5'-CCATGGGGAACGATTCTGTCAGCTACG-3' (SEQ.ID.NO.:45; sense) and  
10   5'-GCTATGCCTGAAGCCAGTCTTGTG-3' (SEQ.ID.NO.:46; antisense)  
and human leukocyte Marathon-Ready cDNA (Clontech) as a template. Advantage cDNA polymerase (Clontech) was used for the amplification in a 50µl reaction by the following cycle with step 2 to step 4 repeated 35 times: 94°C for 30 seconds; 94°C for 10 seconds; 62°C for 20 seconds; 72°C for 1.5 minutes; 72°C for 7 minutes. A 1.0  
15   Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem). The nucleic acid sequence of the novel human receptor RUP10 is set forth in SEQ.ID.NO.:5 and the putative amino acid sequence thereof is set forth in SEQ.ID.NO.:6.

20

**d. hRUP11 (Seq. Id. Nos. 7 & 8)**

          The disclosed human RUP11 was identified based upon the use of GenBank database information. While searching the database, a cDNA clone with accession  
25   number AC013396 was identified as a human genomic sequence from chromosome 2.

The full length RUP11 was cloned by PCR using primers:

5'-CCAGGATGTTGTGTCACCGTGGTGGC-3' (SEQ.ID.NO.:47; sense),

5'-CACAGCGCTGCAGCCCTGCAGCTGGC-3' (SEQ.ID.NO.:48; antisense)

and human genomic DNA (Clontech) as a template. TaqPlus Precision DNA

5 polymerase (Stratagene) was used for the amplification in a 50µl reaction by the following cycle with step 2 to step 4 repeated 35 times: 94°C for 3 minutes; 94°C for 20 seconds; 67°C for 20 seconds; 72°C for 1.5 minutes; 72°C for 7 minutes. A 1.3 Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem). The  
10 nucleic acid sequence of the novel human receptor RUP11 is set forth in SEQ.ID.NO.:7 and the putative amino acid sequence thereof is set forth in SEQ.ID.NO.:8.

**e. hRUP12 (Seq. Id. Nos. 9 & 10)**

The disclosed human RUP12 was identified based upon the use of GenBank database. While searching the database, a cDNA clone with accession number  
15 AP000808 was identified to encode a new GPCR, having significant homology with rat RTA and human mas1 oncogene GPCRs. The full length RUP12 was cloned by PCR using primers:

5'-CTTCCTCTCGTAGGGATGAACCAGAC-3' (SEQ.ID.NO.:49; sense)

5'-CTCGCACAGGTGGGAAGCACCTGTGG-3' (SEQ.ID.NO.:50; antisense)

20 and human genomic DNA (Clontech) as template. TaqPlus Precision DNA polymerase (Stratagene) was used for the amplification by the following cycle with step 2 to step 4 repeated 35 times: 94°C for 3 min; 94°C for 20 sec; 65°C for 20sec; 72°C for 2 min and 72°C for 7 min. A 1.0kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit

(P.E. Biosystem) (*see*, SEQ.ID.NO.:9 for nucleic acid sequence and SEQ.ID.NO.:10 for deduced amino acid sequence).

**f. hRUP13 (Seq. Id. Nos. 11 & 12)**

The disclosed human RUP13 was identified based upon the use of  
5 GenBank database. While searching the database, a cDNA clone with accession number  
AC011780 was identified to encode a new GPCR, having significant homology with  
GPCR fish GPRX-ORYLA. The full length RUP13 was cloned by PCR using primers:  
5'-GCCTGTGACAGGAGGTACCCTGG-3' (SEQ.ID.NO.:51; sense)  
5'-CATATCCCTCCGAGTGTCCAGCGGC-3' (SEQ.ID.NO.:52; antisense)  
10 and human genomic DNA (Clontech) as template. TaqPlus Precision DNA polymerase  
(Stratagene) was used for the amplification by the following cycle with step 2 to step 4  
repeated 35 times: 94°C for 3 min; 94°C for 20 sec; 65°C for 20sec; 72°C for 2 min and  
72°C for 7 min. A 1.35kb PCR fragment was isolated and cloned into the pCRII-TOPO  
vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit  
15 (P.E. Biosystem) (*see*, SEQ.ID.NO.:11 for nucleic acid sequence and SEQ.ID.NO.:12  
for deduced amino acid sequence).

**g. hRUP14 (Seq. Id. Nos. 13 & 14)**

The disclosed human RUP14 was identified based upon the use of GeneBank  
database information. While searching the database, a cDNA clone with Accession  
20 Number AL137118 was identified as a human genomic sequence from chromosome  
13. The full length RUP14 was cloned by PCR using primers:  
5'-GCATGGAGAGAAAATTTATGTCCTTGCAACC-3' (SEQ.ID.NO.:53; sense)  
5'-CAAGAACAGGTCTCATCTAAGAGCTCC-3' (SEQ.ID.NO.:54; antisense)  
and human genomic DNA (Promega) as a template. Taq Plus Precision polymerase  
25 (Stratagene) and 5% DMSO were used for the amplification by the following cycle

with step 2 and step 3 repeated 35 times: 94°C for 3 minute; 94°C for 20 seconds; 58°C for 2 minutes; 72°C for 10 minutes.

A 1.1 Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem) (see, SEQ.ID.NO.:13 for nucleic acid sequence and SEQ.ID.NO.:14 for deduced amino acid sequence). The sequence of RUP14 clones isolated from human genomic DNA matched with the sequence obtained from database.

**h. hRUP15 (Seq. Id. Nos. 15 & 16)**

The disclosed human RUP15 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC016468 was identified as a human genomic sequence. The full length RUP15 was cloned by PCR using primers:

5'-GCTGTTGCCATGACGTCCACCTGCAC-3' (SEQ.ID.NO.:55; sense)

5'-GGACAGTTCAAGGTTTGCCTTAGAAC-3' (SEQ.ID.NO.:56; antisense)

and human genomic DNA (Promega) as a template. Taq Plus Precision polymerase (Stratagene) was used for the amplification by the following cycle with step 2 to 4 repeated 35 times: 94°C for 3 minute; 94°C for 20 seconds; 65°C for 20 seconds; 72°C for 2 minutes and 72°C for 7 minutes.

A 1.5 Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem). See, SEQ.ID.NO.:15 for nucleic acid sequence and SEQ.ID.NO.:16 for deduced amino acid sequence. The sequence of RUP15 clones isolated from human genomic DNA matched with the sequence obtained from database.

**i. hRUP16 (Seq. Id. Nos. 17 & 18)**



The disclosed human RUP16 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AL136106 was identified as a human genomic sequence from chromosome 13. The full length RUP16 was cloned by PCR using primers:

- 5 5'-CTTTCGATACTGCTCCTATGCTC-3' (SEQ.ID.NO.:57; sense, 5' of initiation codon),  
5'-GTAGTCCACTGAAAGTCCAGTGATCC-3' (SEQ.ID.NO.:58; antisense, 3' of stop codon)  
and human skeletal muscle Marathon-Ready cDNA (Clontech) as template. Advantage  
cDNA polymerase (Clontech) was used for the amplification in a 50ul reaction by the  
following cycle with step 2 to 4 repeated 35 times: 94°C for 30 seconds; 94°C for 5  
10 seconds; 69°C for 15 seconds; 72°C for 1 minute and 72°C for 5 minutes.

A 1.1 Kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the T7 sequenase kit (Amsham). See, SEQ.ID.NO.:17 for nucleic acid sequence and SEQ.ID.NO.:18 for deduced amino acid sequence. The sequence of RUP16 clones matched with four unordered segments of  
15 AL136106, indicating that the RUP16 cDNA is composed of 4 exons.

**j. hRUP17 (Seq. Id. Nos. 19 & 20)**

The disclosed human RUP17 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC023078 was identified as a human genomic sequence from chromosome

- 20 11. The full length RUP17 was cloned by PCR using primers:

5'-TTTCTGAGCATGGATCCAACCATCTC-3' (SEQ.ID.NO.:59; sense, containing initiation codon)

5'-CTGTCTGACAGGGCAGAGGCTCTTC-3' (SEQ.ID.NO.:60; antisense, 3' of stop codon)

- and human genomic DNA (Promega) as template. Advantage cDNA polymerase mix  
25 (Clontech) was used for the amplification in a 100ul reaction with 5% DMSO by the

following cycle with step 2 to 4 repeated 30 times: 94°C for 1 min; 94°C for 15 sec;  
67°C for 20 sec; 72°C for 1 min and 30 sec; and 72°C for 5 min.

A 970bp PCR fragment was isolated from 1% agarose gel and cloned into the  
pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye  
5 Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:19 for nucleic acid sequence and  
SEQ.ID.NO.:20 for deduced amino acid sequence.

**k. hRUP18 (Seq. Id. Nos. 21 & 22)**

The disclosed human RUP18 was identified based upon the use of GeneBank  
database information. While searching the database, a cDNA clone with Accession  
10 Number AC008547 was identified as a human genomic sequence from chromosome  
5. The full length RUP18 was cloned by PCR using primers:

5'-GGAAGTCGTATAGACCCAGCGTCGCTCC-3' (SEQ.ID.NO.:61; sense, 5' of the  
initiation codon),

5'-GGAGGTTGCGCCTTAGCGACAGATGACC-3' (SEQ.ID.NO.:62; antisense, 3' of stop  
15 codon)

and human genomic DNA (Promega) as template. TaqPlus precision DNA  
polymerase (Stratagene) was used for the amplification in a 100ul reaction with 5%  
DMSO by the following cycle with step 2 to 4 repeated 35 times: 95°C for 5 min;  
95°C for 30 sec; 65°C for 30 sec; 72°C for 2 min; and 72°C for 5 min.

20 A 1.3kb PCR fragment was isolated from 1% agarose gel and cloned into the  
pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye  
Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:21 for nucleic acid sequence and  
SEQ.ID.NO.:22 for deduced amino acid sequence.

**l. hRUP19 (Seq. Id. Nos. 23 & 24)**

The disclosed human RUP19 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC026331 was identified as a human genomic sequence from chromosome 12. The full length RUP19 was cloned by PCR using primers:

- 5 5'-CTGCACCCGGACACTTGCTCTG-3' (SEQ.ID.NO.:63; sense, 5' of initiation codon),  
 5'-GTCTGCTTGTTCAGTGCCACTCAAC-3' (SEQ.ID.NO.:64; antisense, containing the stop  
 codon)  
 and human genomic DNA (Promega) as template. TaqPlus Precision DNA  
 polymerase (Stratagene) was used for the amplification with 5% DMSO by the  
 10 following cycle with step 2 to 4 repeated 35 times: 94°C for 1 min; 94°C for 15 sec;  
 70°C for 20 sec; 72°C for 1 min and 30 sec; and 72°C for 5 min.

- A 1.1kp PCR fragment was isolated from 1% agarose gel and cloned into the  
 pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye  
 Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:23 for nucleic acid sequence and  
 15 SEQ.ID.NO.:24 for deduced amino acid sequence.

#### m. hRUP20 (Seq. Id. Nos. 25 & 26)

- The disclosed human RUP20 was identified based upon the use of GeneBank  
 database information. While searching the database, a cDNA clone with Accession  
 Number AL161458 was identified as a human genomic sequence from chromosome  
 20 1. The full length RUP20 was cloned by PCR using primers:  
 5'-TATCTGCAATTCTATTCTAGCTCCTG-3' (SEQ.ID.NO.:65; sense, 5' of initiation codon),  
 5'-TGTCCCTAATAAAGTCACATGAATGC-3' (SEQ.ID.NO.:66; antisense, 3' of stop codon)  
 and human genomic DNA (Promega) as template. Advantage cDNA polymerase mix  
 (Clonotech) was used for the amplification with 5% DMSO by the following cycle with

step 2 to 4 repeated 35 times: 94°C for 1 min; 94°C for 15 sec; 60°C for 20 sec; 72°C for 1 min and 30 sec; and 72°C for 5 min.

A 1.0 kp PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye  
5 Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:25 for nucleic acid sequence and SEQ.ID.NO.:26 for deduced amino acid sequence.

**n. hRUP21 (Seq. Id. Nos. 27 & 28)**

The disclosed human RUP21 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession  
10 Number AC026756 was identified as a human genomic sequence from chromosome 13. The full length RUP21 was cloned by PCR using primers:

5'-GGAGACAACCATGAATGAGCCAC-3' (SEQ.ID.NO.:67; sense)

5'-TATTCAAGGGTTGTTTGAGTAAC-3' (SEQ.ID.NO.:68; antisense)

and human genomic DNA (Promega) as template. Taq Plus Precision polymerase  
15 (Stratagene) was used for the amplification in a 100ul reaction with 5% DMSO by the following cycle with step 2 to 4 repeated 30 times: 94°C for 1 min; 94°C for 15 sec; 55°C for 20 sec; 72°C for 1 min and 30 sec; and 72°C for 5 min.

A 1,014 bp PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye  
20 Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:27 for nucleic acid sequence and SEQ.ID.NO.:28 for deduced amino acid sequence.

**o. hRUP22 (Seq. Id. Nos. 29 & 30)**

The disclosed human RUP22 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession

Number AC027026 was identified as a human genomic sequence from chromosome

11. The full length RUP22 was cloned by PCR using primers:

5'-GGCACCAGTGGAGGTTTTCTGAGCATG-3' (SEQ.ID.NO.:69; sense, containing initiation codon)

5 5'-CTGATGGAAGTAGAGGCTGTCCATCTC-3' (SEQ.ID.NO.:70; antisense, 3' of stop codon)

and human genomic DNA (Promega) as template. TaqPlus Precision DNA polymerase (Stratagene) was used for the amplification in a 100ul reaction with 5% DMSO by the following cycle with step 2 to 4 repeated 30 times: 94°C, 1 minutes 94°C, 15 seconds

10 55°C, 20 seconds 72°C, 1.5 minute 72°C, 5 minutes.

A 970bp PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator Kit (P.E. Biosystem). See, SEQ.ID.NO.:29 for nucleic acid sequence and SEQ.ID.NO.:30 for deduced amino acid sequence.

15 **p. hRUP23 (Seq. Id. Nos. 31 & 32)**

The disclosed human RUP23 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC007104 was identified as a human genomic sequence from chromosome 4. The full length RUP23 was cloned by PCR using primers:

20 5'-CCTGGCGAGCCGCTAGCGCCATG-3' (SEQ.ID.NO.:71; sense, ATG as the initiation codon),

5'-ATGAGCCCTGCCAGGCCCTCAGT-3' (SEQ.ID.NO.:72; antisense, TCA as the stop codon)

and human placenta Marathon-Ready cDNA (Clontech) as template. Advantage cDNA

25 polymerase (Clontech) was used for the amplification in a 50ul reaction by the following

cycle with step 2 to 4 repeated 35 times: 95°C for 30 sec; 95°C for 15 sec; 66°C for 20 sec; 72°C for 1 min and 20 sec; and 72°C for 5 min.

A 1.0 kb PCR fragment was isolated and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator Kit (P.E. Biosystem). See, SEQ.ID.NO.:31 for nucleic acid sequence and SEQ.ID.NO.:32 for deduced amino acid sequence.

**q. hRUP24 (Seq. Id. Nos. 33 & 34)**

The disclosed human RUP25 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC026331 was identified as a human genomic sequence from chromosome 12. The full length RUP25 was cloned by PCR using primers:  
5'-GCTGGAGCATTCACTAGGCGAG-3' (SEQ.ID.NO.:73; sense, 5' of initiation codon),  
5'-AGATCCTGGTTCTTGGTGACAATG-3' (SEQ.ID.NO.:74; antisense, 3' of stop codon)  
and human genomic DNA (Promega) as template. Advantage cDNA polymerase mix (Clontech) was used for the amplification with 5% DMSO by the following cycle with step 2 to 4 repeated 35 times: 94°C for 1 minute; 94°C for 15 seconds; 56°C for 20 seconds 72°C for 1 minute 30 seconds and 72°C for 5 minutes.

A 1.2kb PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator Kit (P.E. Biosystem). See, SEQ.ID.NO.:33 for nucleic acid sequence and SEQ.ID.NO.:34 for deduced amino acid sequence.

**r. hRUP25 (Seq. Id. Nos. 35 & 36)**

The disclosed human RUP25 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession

Number AC026331 was identified as a human genomic sequence from chromosome

12. The full length RUP25 was cloned by PCR using primers:

5'-GCTGGAGCATTCAGGCGAG-3' (SEQ.ID.NO.:75; sense, 5' of initiation codon),

5'-AGATCCTGGTTCTTGGTGACAATG-3' (SEQ.ID.NO.:76; antisense, 3' of stop codon)

5 and human genomic DNA (Promega) as template. Advantage cDNA polymerase mix (Clontech) was used for the amplification with 5% DMSO by the following cycle with step 2 to 4 repeated 35 times: 94°C for 1 minute; 94°C for 15 seconds; 56°C for 20 seconds 72°C for 1 minute 30 seconds and 72°C for 5 minutes.

A 1.2kb PCR fragment was isolated from 1% agarose gel and cloned into the  
10 pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Terminator Kit (P.E. Biosystem). See, SEQ.ID.NO.:35 for nucleic acid sequence and SEQ.ID.NO.:36 for deduced amino acid sequence.

**s. hRUP26 (Seq. Id. Nos. 37 & 38)**

The disclosed human RUP26 was identified based upon the use of GeneBank  
15 database information. While searching the database, a cDNA clone with Accession Number AC023040 was identified as a human genomic sequence from chromosome 2. The full length RUP26 was cloned by RT-PCR using RUP26 specific primers:

5'-AGCCATCCCTGCCAGGAAGCATGG-3' (SEQ.ID.NO.:77; sense, containing initiation codon)

20 5'-CCAGACTGTGGACTCAAGAACTCTAGG-3' (SEQ.ID.NO.:78; antisense, containing stop codon)

and human pancreas Marathon - Ready cDNA (Clontech) as template. Advantage cDNA polymerase mix (Clontech) was used for the amplification in a 100µl reaction with 5% DMSO by the following cycle with step 2 to 4 repeated 35 times: 94°C for 5 minute;  
25 95°C for 30 seconds; 65°C for 30 seconds 72°C for 2 minute and 72°C for 5 minutes.

A 1.1kb PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:37 for nucleic acid sequence and SEQ.ID.NO.:38 for deduced amino acid sequence.

5                    **t. hRUP27 (Seq. Id. Nos. 39 & 40)**

The disclosed human RUP27 was identified based upon the use of GeneBank database information. While searching the database, a cDNA clone with Accession Number AC027643 was identified as a human genomic sequence from chromosome 12. The full length RUP27 was cloned by PCR using RUP27 specific primers:

- 10    5'-AGTCCACGAACAATGAATCCATTTCATG-3' (SEQ.ID.NO.:79; sense, containing initiation codon),
- 5'-ATCATGTCTAGACTCATGGTGATCC-3' (SEQ.ID.NO.:80; antisense, 3' of stop codon)
- and the human adult brain Marathon-Ready cDNA (Clontech) as template. Advantage cDNA polymerase mix (Clontech) was used for the amplification in a 50µl reaction with 5%
- 15    DMSO by the following cycle with step 2 to 4 repeated 35 times: 94°C for 1 minute; 94°C for 10 seconds; 58°C for 20 seconds 72°C for 1 minute 30 seconds and 72°C for 5 minutes.

- A 1.1kb PCR fragment was isolated from 1% agarose gel and cloned into the pCRII-TOPO vector (Invitrogen) and completely sequenced using the ABI Big Dye
- 20    Termiantor Kit (P.E. Biosystem). See, SEQ.ID.NO.:35 for nucleic acid sequence and SEQ.ID.NO.:36 for deduced amino acid sequence. The sequence of RUP27 cDNA clone isolated from human brain was determined to match with five unordered segments of AC027643, indicating that the RUP27 cDNA is composed of 5 exons.



**Example 2****PREPARATION OF NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED GPCRS**

Those skilled in the art are credited with the ability to select techniques for mutation of a nucleic acid sequence. Presented below are approaches utilized to create non-endogenous versions of several of the human GPCRs disclosed above. The mutations disclosed below are based upon an algorithmic approach whereby the 16<sup>th</sup> amino acid (located in the IC3 region of the GPCR) from a conserved proline (or an endogenous, conservative substitution therefore) residue (located in the TM6 region of the GPCR, near the TM6/IC3 interface) is mutated, preferably to an alanine, histidine, arginine or lysine amino acid residue, most preferably to a lysine amino acid residue.

**1. Transformer Site-Directed™ Mutagenesis**

Preparation of non-endogenous human GPCRs may be accomplished on human GPCRs using Transformer Site-Directed™ Mutagenesis Kit (Clontech) according to the manufacturer instructions. Two mutagenesis primers are utilized, most preferably a lysine mutagenesis oligonucleotide that creates the lysine mutation, and a selection marker oligonucleotide. For convenience, the codon mutation to be incorporated into the human GPCR is also noted, in standard form (Table D):

**TABLE D**

Receptor Identifier	Codon Mutation
<b>hRUP8</b>	V274K
<b>hRUP9</b>	T249K
<b>hRUP10</b>	R232K
<b>hRUP11</b>	M294K
<b>hRUP12</b>	F220K
<b>hRUP16</b>	A238K

hRUP17	Y215K
hRUP18	L294K
hRUP19	T219K
hRUP20	K248A K248H K248R
hRUP21	R240K
hRUP22	Y222K
hRUP24	A245K
hRUP25	I230K
hRUP26	V285K
hRUP27	T248K

## 2. QuikChange™ Site-Directed™ Mutagenesis

Preparation of non-endogenous human GPCRs can also be accomplished by using QuikChange™ Site-Directed™ Mutagenesis Kit (Stratagene, according to manufacturer's instructions). Endogenous GPCR is preferably used as a template and two mutagenesis primers utilized, as well as, most preferably, a lysine mutagenesis oligonucleotide and a selection marker oligonucleotide (included in kit). For convenience, the codon mutation incorporated into the novel human GPCR and the respective oligonucleotides are noted, in standard form (Table E):

**TABLE E**

Receptor Identifier	Codon Mutation	5'-3' orientation (sense), (SEQ.ID.NO.) mutation underlined	5'-3' orientation (antisense) (SEQ.ID.NO.)	Cycle Conditions Min ('), Sec (") Cycles 2-4 repeated 16 times
hRUP13	A268K	GGGGAGGGAAAGCAA <u>AGGTGGTCCTCCTGG</u> (81)	CCAGGAGAACCACCT <u>TTGCTTTCCCTCCCC</u> (82)	98° for 2' 98° for 30" 56°C for 30" 72° for 11' 40" 72° for 5'
hRUP14	L246K	CAGGAAGGCAAAGAC <u>CACCATCATCATC</u> (85)	GATGATGATGGTGGT <u>CTTGCCTTCCTG</u> (86)	98° for 2' 98° for 30" 55°C for 30" 72° for 11' 40" 72° for 5'

hRUP15	A398K	CCAGTGCAAAGCTAAG AAAGTGATCTTC (89)	GAAGATCACTTTCTTA GCTTTGCACTGG (90)	98° for 2' 98° for 30" 55°C for 30" 72° for 11' 40" 72° for 5'
hRUP23	W275K	GCCGCCACCGCGCCAA GAGGAAGATTGGC (93)	GCCAATCTTCCTCTTG GCGCGGTGGCGGC (94)	98° for 2' 98° for 30" 56°C for 30" 72° for 11' 40" 72° for 5'

The non-endogenous human GPCRs were then sequenced and the derived and verified nucleic acid and amino acid sequences are listed in the accompanying "Sequence Listing" appendix to this patent document, as summarized in Table F below:

TABLE F

Non Endogenous Human GPCR	Nucleic Acid Sequence Listing	Amino Acid Sequence Listing
hRUP13	SEQ.ID.NO.:83	SEQ.ID.NO.:84
hRUP14	SEQ.ID.NO.:87	SEQ.ID.NO.:88
hRUP15	SEQ.ID.NO.:91	SEQ.ID.NO.:92
hRUP23	SEQ.ID.NO.:95	SEQ.ID.NO.:96

### Example 3

#### 10 RECEPTOR EXPRESSION

Although a variety of cells are available to the art for the expression of proteins, it is most preferred that mammalian cells be utilized. The primary reason for this is predicated upon practicalities, *i.e.*, utilization of, *e.g.*, yeast cells for the expression of a GPCR, while possible, introduces into the protocol a non-mammalian cell which may not (indeed, in the case of yeast, does not) include the receptor-coupling, genetic-mechanism and secretory pathways that have evolved for mammalian systems – thus, results obtained in non-mammalian cells, while of

potential use, are not as preferred as that obtained from mammalian cells. Of the mammalian cells, COS-7, 293 and 293T cells are particularly preferred, although the specific mammalian cell utilized can be predicated upon the particular needs of the artisan.

5           **a.       Transient Transfection**

On day one,  $6 \times 10^6$  10 cm dish of 293 cells well were plated out. On day two, two reaction tubes were prepared (the proportions to follow for each tube are per plate): tube A was prepared by mixing  $4 \mu\text{g}$  DNA (*e.g.*, pCMV vector; pCMV vector with receptor cDNA, etc.) in 0.5 ml serum free DMEM (Gibco BRL); tube B was prepared by  
10 mixing  $24 \mu\text{l}$  lipofectamine (Gibco BRL) in 0.5ml serum free DMEM. Tubes A and B were admixed by inversions (several times), followed by incubation at room temperature for 30–45min. The admixture is referred to as the “transfection mixture”. Plated 293 cells were washed with 1XPBS, followed by addition of 5 ml serum free DMEM. 1 ml of the transfection mixture were added to the cells, followed by incubation for 4hrs at  
15  $37^\circ\text{C}/5\% \text{CO}_2$ . The transfection mixture was removed by aspiration, followed by the addition of 10ml of DMEM/10% Fetal Bovine Serum. Cells were incubated at  $37^\circ\text{C}/5\% \text{CO}_2$ . After 48hr incubation, cells were harvested and utilized for analysis.

**b.       Stable Cell Lines: Gs Fusion Protein**

Approximately  $12 \times 10^6$  293 cells are plated on a 15cm tissue culture plate.  
20 Grown in DME High Glucose Medium containing ten percent fetal bovine serum and one percent sodium pyruvate, L-glutamine, and anti-biotics. Twenty-four hours following plating of 293 cells to ~80% confluency, the cells are transfected using  $12 \mu\text{g}$  of DNA. The  $12 \mu\text{g}$  of DNA is combined with 60ul of lipofectamine and 2mL of DME High Glucose Medium without serum. The medium is aspirated from the plates and the  
25 cells are washed once with medium without serum. The DNA, lipofectamine, and

medium mixture is added to the plate along with 10mL of medium without serum. Following incubation at 37 degrees Celsius for four to five hours, the medium is aspirated and 25ml of medium containing serum is added. Twenty-four hours following transfection, the medium is aspirated again, and fresh medium with serum is added.

5 Forty-eight hours following transfection, the medium is aspirated and medium with serum is added containing geneticin (G418 drug) at a final concentration of 500µg/mL. The transfected cells now undergo selection for positively transfected cells containing the G418 resistant gene. The medium is replaced every four to five days as selection occurs. During selection, cells are grown to create stable pools, or split for stable clonal

10 selection.

#### **Example 4**

#### **ASSAYS FOR DETERMINATION OF CONSTITUTIVE ACTIVITY OF NON-ENDOGENOUS GPCRS**

15 A variety of approaches are available for assessment of constitutive activity of the non-endogenous human GPCRs. The following are illustrative; those of ordinary skill in the art are credited with the ability to determine those techniques that are preferentially beneficial for the needs of the artisan.

#### **1. Membrane Binding Assays: [<sup>35</sup>S]GTPγS Assay**

20 When a G protein-coupled receptor is in its active state, either as a result of ligand binding or constitutive activation, the receptor couples to a G protein and stimulates the release of GDP and subsequent binding of GTP to the G protein. The alpha subunit of the G protein-receptor complex acts as a GTPase and slowly hydrolyzes

25 the GTP to GDP, at which point the receptor normally is deactivated. Constitutively activated receptors continue to exchange GDP for GTP. The non-hydrolyzable GTP analog, [<sup>35</sup>S]GTPγS, can be utilized to demonstrate enhanced binding of [<sup>35</sup>S]GTPγS to membranes expressing constitutively activated receptors. The advantage of using

[<sup>35</sup>S]GTPγS binding to measure constitutive activation is that: (a) it is generically applicable to all G protein-coupled receptors; (b) it is proximal at the membrane surface making it less likely to pick-up molecules which affect the intracellular cascade.

The assay utilizes the ability of G protein coupled receptors to stimulate  
5 [<sup>35</sup>S]GTPγS binding to membranes expressing the relevant receptors. The assay can, therefore, be used in the direct identification method to screen candidate compounds to known, orphan and constitutively activated G protein-coupled receptors. The assay is generic and has application to drug discovery at all G protein-coupled receptors.

The [<sup>35</sup>S]GTPγS assay was incubated in 20 mM HEPES and between 1 and  
10 about 20mM MgCl<sub>2</sub> (this amount can be adjusted for optimization of results, although 20mM is preferred) pH 7.4, binding buffer with between about 0.3 and about 1.2 nM [<sup>35</sup>S]GTPγS (this amount can be adjusted for optimization of results, although 1.2 is preferred ) and 12.5 to 75 μg membrane protein (*e.g.* 293 cells expressing the Gs Fusion Protein; this amount can be adjusted for optimization) and 10 μM GDP (this amount can  
15 be changed for optimization) for 1 hour. Wheatgerm agglutinin beads (25 μl; Amersham) were then added and the mixture incubated for another 30 minutes at room temperature. The tubes were then centrifuged at 1500 x g for 5 minutes at room temperature and then counted in a scintillation counter.

## 2. Adenylyl Cyclase

20 A Flash Plate™ Adenylyl Cyclase kit (New England Nuclear, Cat. No. SMP004A) designed for cell-based assays can be modified for use with crude plasma membranes. The Flash Plate wells can contain a scintillant coating which also contains a specific antibody recognizing cAMP. The cAMP generated in the wells can be quantitated by a direct competition for binding of radioactive cAMP tracer to the cAMP

antibody. The following serves as a brief protocol for the measurement of changes in cAMP levels in whole cells that express the receptors.

Transfected cells were harvested approximately twenty four hours after transient transfection. Media is carefully aspirated off and discarded. 10ml of PBS is gently  
5 added to each dish of cells followed by careful aspiration. 1ml of Sigma cell dissociation buffer and 3ml of PBS are added to each plate. Cells were pipeted off the plate and the cell suspension was collected into a 50ml conical centrifuge tube. Cells were then centrifuged at room temperature at 1,100 rpm for 5 min. The cell pellet was carefully re-suspended into an appropriate volume of PBS (about 3ml/plate). The cells  
10 were then counted using a hemocytometer and additional PBS was added to give the appropriate number of cells (with a final volume of about 50  $\mu$ l/well).

cAMP standards and Detection Buffer (comprising 1  $\mu$ Ci of tracer [ $^{125}$ I] cAMP (50  $\mu$ l] to 11 ml Detection Buffer) was prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer was prepared fresh for screening and  
15 contained 50 $\mu$ l of Stimulation Buffer, 3 $\mu$ l of test compound (12 $\mu$ M final assay concentration) and 50 $\mu$ l cells, Assay Buffer was stored on ice until utilized. The assay was initiated by addition of 50 $\mu$ l of cAMP standards to appropriate wells followed by addition of 50 $\mu$ l of PBSA to wells H-11 and H12. 50 $\mu$ l of Stimulation Buffer was added to all wells. DMSO (or selected candidate compounds) was added to appropriate wells  
20 using a pin tool capable of dispensing 3 $\mu$ l of compound solution, with a final assay concentration of 12 $\mu$ M test compound and 100 $\mu$ l total assay volume. The cells were then added to the wells and incubated for 60 min at room temperature. 100 $\mu$ l of Detection Mix containing tracer cAMP was then added to the wells. Plates were then incubated additional 2 hours followed by counting in a Wallac MicroBeta scintillation

counter. Values of cAMP/well were then extrapolated from a standard cAMP curve which was contained within each assay plate.

### 3. Cell-Based cAMP for Gi Coupled Target GPCRs

5 TSHR is a Gs coupled GPCR that causes the accumulation of cAMP upon activation. TSHR will be constitutively activated by mutating amino acid residue 623 (*i.e.*, changing an alanine residue to an isoleucine residue). A Gi coupled receptor is expected to inhibit adenylyl cyclase, and, therefore, decrease the level of cAMP production, which can make assessment of cAMP levels challenging. An effective  
10 technique for measuring the decrease in production of cAMP as an indication of constitutive activation of a Gi coupled receptor can be accomplished by co-transfecting, most preferably, non-endogenous, constitutively activated TSHR (TSHR-A623I) (or an endogenous, constitutively active Gs coupled receptor) as a "signal enhancer" with a Gi linked target GPCR to establish a baseline level of cAMP. Upon creating a non-  
15 endogenous version of the Gi coupled receptor, this non-endogenous version of the target GPCR is then co-transfected with the signal enhancer, and it is this material that can be used for screening. We will utilize such approach to effectively generate a signal when a cAMP assay is used; this approach is preferably used in the direct identification of candidate compounds against Gi coupled receptors. It is noted that for a Gi coupled  
20 GPCR, when this approach is used, an inverse agonist of the target GPCR will increase the cAMP signal and an agonist will decrease the cAMP signal.

On day one,  $2 \times 10^4$  293 and 293 cells/well will be plated out. On day two, two reaction tubes will be prepared (the proportions to follow for each tube are per plate): tube A will be prepared by mixing 2 $\mu$ g DNA of each receptor transfected into the  
25 mammalian cells, for a total of 4 $\mu$ g DNA (*e.g.*, pCMV vector, pCMV vector with mutated THSR (TSHR-A623I); TSHR-A623I and GPCR, etc.) in 1.2ml serum free



DMEM (Irvine Scientific, Irvine, CA); tube B will be prepared by mixing 120 $\mu$ l lipofectamine (Gibco BRL) in 1.2ml serum free DMEM. Tubes A and B will then be admixed by inversions (several times), followed by incubation at room temperature for 30-45min. The admixture is referred to as the "transfection mixture". Plated 293 cells  
5 will be washed with 1XPBS, followed by addition of 10ml serum free DMEM. 2.4ml of the transfection mixture will then be added to the cells, followed by incubation for 4hrs at 37°C/5% CO<sub>2</sub>. The transfection mixture will then be removed by aspiration, followed by the addition of 25ml of DMEM/10% Fetal Bovine Serum. Cells will then be incubated at 37°C/5% CO<sub>2</sub>. After 24hr incubation, cells will then be harvested and  
10 utilized for analysis.

A Flash Plate™ Adenylyl Cyclase kit (New England Nuclear; Cat. No. SMP004A) is designed for cell-based assays, however, can be modified for use with crude plasma membranes depending on the need of the skilled artisan. The Flash Plate wells will contain a scintillant coating which also contains a specific antibody  
15 recognizing cAMP. The cAMP generated in the wells can be quantitated by a direct competition for binding of radioactive cAMP tracer to the cAMP antibody. The following serves as a brief protocol for the measurement of changes in cAMP levels in whole cells that express the receptors.

Transfected cells will be harvested approximately twenty four hours after  
20 transient transfection. Media will be carefully aspirated off and discarded. 10ml of PBS will be gently added to each dish of cells followed by careful aspiration. 1ml of Sigma cell dissociation buffer and 3ml of PBS will be added to each plate. Cells will be pipetted off the plate and the cell suspension will be collected into a 50ml conical centrifuge tube. Cells will then be centrifuged at room temperature at 1,100 rpm for 5 min. The cell  
25 pellet will be carefully re-suspended into an appropriate volume of PBS (about

3ml/plate). The cells will then be counted using a hemocytometer and additional PBS is added to give the appropriate number of cells (with a final volume of about 50 $\mu$ l/well).

cAMP standards and Detection Buffer (comprising 1  $\mu$ Ci of tracer [ $^{125}$ I] cAMP (50  $\mu$ l] to 11 ml Detection Buffer) will be prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer should be prepared fresh for screening and contained 50 $\mu$ l of Stimulation Buffer, 3 $\mu$ l of test compound (12 $\mu$ M final assay concentration) and 50 $\mu$ l cells, Assay Buffer can be stored on ice until utilized. The assay can be initiated by addition of 50 $\mu$ l of cAMP standards to appropriate wells followed by addition of 50 $\mu$ l of PBSA to wells H-11 and H12. 50 $\mu$ l of Stimulation Buffer will be added to all wells. Selected compounds (e.g., TSH) will be added to appropriate wells using a pin tool capable of dispensing 3 $\mu$ l of compound solution, with a final assay concentration of 12 $\mu$ M test compound and 100 $\mu$ l total assay volume. The cells will then be added to the wells and incubated for 60 min at room temperature. 100 $\mu$ l of Detection Mix containing tracer cAMP will then be added to the wells. Plates were then incubated additional 2 hours followed by counting in a Wallac MicroBeta scintillation counter. Values of cAMP/well will then be extrapolated from a standard cAMP curve which is contained within each assay plate.

#### 4. Reporter-Based Assays

##### a. CRE-LUC Reporter Assay (Gs-associated receptors)

293 and 293T cells are plated-out on 96 well plates at a density of  $2 \times 10^4$  cells per well and were transfected using Lipofectamine Reagent (BRL) the following day according to manufacturer instructions. A DNA/lipid mixture is prepared for each 6-well transfection as follows: 260ng of plasmid DNA in 100 $\mu$ l of DMEM were gently mixed with 2 $\mu$ l of lipid in 100 $\mu$ l of DMEM (the 260ng of plasmid DNA consisted of

200ng of a 8xCRE-Luc reporter plasmid, 50ng of pCMV comprising endogenous receptor or non-endogenous receptor or pCMV alone, and 10ng of a GPRS expression plasmid (GPRS in pcDNA3 (Invitrogen)). The 8XCRE-Luc reporter plasmid was prepared as follows: vector SRIF- $\beta$ -gal was obtained by cloning the rat somatostatin promoter (-71/+51) at BglV-HindIII site in the p $\beta$ gal-Basic Vector (Clontech). Eight (8) copies of cAMP response element were obtained by PCR from an adenovirus template AdpCF126CCRE8 (see, *7 Human Gene Therapy* 1883 (1996)) and cloned into the SRIF- $\beta$ -gal vector at the Kpn-BglV site, resulting in the 8xCRE- $\beta$ -gal reporter vector. The 8xCRE-Luc reporter plasmid was generated by replacing the beta-galactosidase gene in the 8xCRE- $\beta$ -gal reporter vector with the luciferase gene obtained from the pGL3-basic vector (Promega) at the HindIII-BamHI site. Following 30 min. incubation at room temperature, the DNA/lipid mixture was diluted with 400  $\mu$ l of DMEM and 100 $\mu$ l of the diluted mixture was added to each well. 100  $\mu$ l of DMEM with 10% FCS were added to each well after a 4hr incubation in a cell culture incubator. The following day the transfected cells were changed with 200  $\mu$ l/well of DMEM with 10% FCS. Eight (8) hours later, the wells were changed to 100  $\mu$ l /well of DMEM without phenol red, after one wash with PBS. Luciferase activity were measured the next day using the LucLite™ reporter gene assay kit (Packard) following manufacturer instructions and read on a 1450 MicroBeta™ scintillation and luminescence counter (Wallac).

**b. AP1 reporter assay (Gq-associated receptors)**

A method to detect Gq stimulation depends on the known property of Gq-dependent phospholipase C to cause the activation of genes containing AP1 elements in their promoter. A Pathdetect™ AP-1 cis-Reporting System (Stratagene, Catalogue # 219073) can be utilized following the protocol set forth above with respect to the

CREB reporter assay, except that the components of the calcium phosphate precipitate were 410 ng pAPI-Luc, 80 ng pCMV-receptor expression plasmid, and 20 ng CMV-SEAP.

**c. SRF-LUC Reporter Assay (Gq- associated receptors)**

5 One method to detect Gq stimulation depends on the known property of Gq-dependent phospholipase C to cause the activation of genes containing serum response factors in their promoter. A Pathdetect™ SRF-Luc-Reporting System (Stratagene) can be utilized to assay for Gq coupled activity in, *e.g.*, COS7 cells. Cells are transfected with the plasmid components of the system and the indicated  
10 expression plasmid encoding endogenous or non-endogenous GPCR using a Mammalian Transfection™ Kit (Stratagene, Catalogue #200285) according to the manufacturer's instructions. Briefly, 410 ng SRF-Luc, 80 ng pCMV-receptor expression plasmid and 20 ng CMV-SEAP (secreted alkaline phosphatase expression plasmid; alkaline phosphatase activity is measured in the media of transfected cells to  
15 control for variations in transfection efficiency between samples) are combined in a calcium phosphate precipitate as per the manufacturer's instructions. Half of the precipitate is equally distributed over 3 wells in a 96-well plate, kept on the cells in a serum free media for 24 hours. The last 5 hours the cells are incubated with 1μM Angiotensin, where indicated. Cells are then lysed and assayed for luciferase activity  
20 using a Lucite™ Kit (Packard, Cat. # 6016911) and "Trilux 1450 Microbeta" liquid scintillation and luminescence counter (Wallac) as per the manufacturer's instructions. The data can be analyzed using GraphPad Prism™ 2.0a (GraphPad Software Inc.).

**d. Intracellular IP<sub>3</sub> Accumulation Assay (Gq-associated receptors)**

On day 1, cells comprising the receptors (endogenous and/or non-endogenous)  
5 can be plated onto 24 well plates, usually  $1 \times 10^5$  cells/well (although this number can be optimized. On day 2 cells can be transfected by firstly mixing 0.25  $\mu$ g DNA in 50  $\mu$ l serum free DMEM/well and 2  $\mu$ l lipofectamine in 50  $\mu$ l serumfree DMEM/well. The solutions are gently mixed and incubated for 15-30 min at room temperature. Cells are washed with 0.5 ml PBS and 400  $\mu$ l of serum free media is mixed with the transfection  
10 media and added to the cells. The cells are then incubated for 3-4 hrs at 37°C/5%CO<sub>2</sub> and then the transfection media is removed and replaced with 1ml/well of regular growth media. On day 3 the cells are labeled with <sup>3</sup>H-myo-inositol. Briefly, the media is removed and the cells are washed with 0.5 ml PBS. Then 0.5 ml inositol-free/serum free media (GIBCO BRL) is added/well with 0.25  $\mu$ Ci of <sup>3</sup>H-myo-inositol/ well and the cells  
15 are incubated for 16-18 hrs o/n at 37°C/5%CO<sub>2</sub>. On Day 4 the cells are washed with 0.5 ml PBS and 0.45 ml of assay medium is added containing inositol-free/serum free media 10  $\mu$ M pargyline 10 mM lithium chloride or 0.4 ml of assay medium and 50  $\mu$ l of 10x ketanserin (ket) to final concentration of 10  $\mu$ M. The cells are then incubated for 30 min at 37°C. The cells are then washed with 0.5 ml PBS and 200  $\mu$ l of fresh/icecold stop  
20 solution (1M KOH; 18 mM Na-borate; 3.8 mM EDTA) is added/well. The solution is kept on ice for 5-10 min or until cells were lysed and then neutralized by 200  $\mu$ l of fresh/ice cold neutralization sol. (7.5 % HCL). The lysate is then transferred into 1.5 ml eppendorf tubes and 1 ml of chloroform/methanol (1:2) is added/tube. The solution is vortexed for 15 sec and the upper phase is applied to a Biorad AG1-X8™ anion  
25 exchange resin (100-200 mesh). Firstly, the resin is washed with water at 1:1.25 W/V and 0.9 ml of upper phase is loaded onto the column. The column is washed with 10 mls of 5 mM myo-inositol and 10 ml of 5 mM Na-borate/60mM Na-formate. The inositol

tris phosphates are eluted into scintillation vials containing 10 ml of scintillation cocktail with 2 ml of 0.1 M formic acid/ 1 M ammonium formate. The columns are regenerated by washing with 10 ml of 0.1 M formic acid/3M ammonium formate and rinsed twice with dd H<sub>2</sub>O and stored at 4°C in water.

5 Exemplary results are presented below in Table G:

**TABLE G**

Receptor	Mutation	Assay Utilized (Figure No.)	Signal Generated: CMV	Signal Generated: Endogenous Version (Relative Light Units)	Signal Generated: Non-Endogenous Version (Relative Light Units)	Difference (=) Between CMV v. Wild-type v. Mutant
hRUP12	N/A	IP <sub>3</sub> (Figure 1)	317.03 cpm/mg protein	3463.29 cpm/mg protein	--	1. 11 Fold =
hRUP13	N/A	cAMP (Figure 2)	8.06 pmol/cAMP/mg protein	19.10 pmol/cAMP/mg protein	--	1. 2.4 Fold =
	A268K	8XCRE- LUC (Figure 3)	3665.43 LCPS	83280.17 LCPS	61713.6 LCPS	1. 23 Fold = 2. 26 % <
hRUP14	L246K	8XCRE- LUC (Figure 5)	86.07 LCPS	1962.87 LCPS	789.73 LCPS	1. 23 Fold = 2. 60% <
hRUP15	A398K	8XCRE- LUC (Figure 6)	86.07 LCPS	18286.77 LCPS	17034.83 LCPS	1. 212 Fold = 2. 1% <
	A398K	cAMP (Figure 7)	15.00 pmol/cAMP/mg protein	164.4 pmol/cAMP/mg protein	117.5 pmol/cAMP/ mg protein	1. 11 Fold = 2. 29% <
hRUP17	N/A	IP <sub>3</sub> (Figure 9)	317.03 cpm/mg protein	741.07 cpm/mg protein	--	1. 2.3 Fold =
hRUP21	N/A	IP <sub>3</sub> (Figure 10)	730.5 cpm/mg protein	1421.9 cpm/mg protein	--	1. 2 Fold =
hRUP23	W275K	8XCRE- LUC (Figure 11)	311.73 pmol/cAMP/mg protein	13756.00 pmol/cAMP/mg protein	9756.87 pmol/cAMP/ mg protein	1. 44 Fold = 2. 30% <

N/A = not applied

Exemplary results of GTP $\gamma$ S assay for detecting constitutive activation, as disclosed in Example 4(1) above, was accomplished utilizing Gs:Fusion Protein Constructs on human RUP13 and RUP15. Table H below lists the signals generated from this assay and the difference in signals as indicated:

5

TABLE H

Receptor: Gs Fusion Protein	Assay Utilized	Signal Generated: CMV (cpm bound GTP)	Signal Generated: Fusion Protein (cpm bound GTP)	Signal Generated: CMV+ 10 $\mu$ M GDP (cpm bound GTP)	Signal Generated: Fusion Protein + 10 $\mu$ M GDP (cpm bound GTP)	Difference Between: 1. CMV v. Fusion Protein 2. CMV+GDP vs. Fusion+GDP 3. Fusion vs. Fusion+GDP (cpm bound GTP)
hRUP13-Gs	GTP $\gamma$ S (Figure 4)	32494.0	49351.30	11148.30	28834.67	1. 1.5 Fold $\Leftarrow$ 2. 2.6 Fold $\Leftarrow$ 3. 42% $\langle$
hRUP15-Gs	GTP $\gamma$ S (Figure 8)	30131.67	32493.67	7697.00	14157.33	1. 1.1 Fold $\Leftarrow$ 2. 1.8 Fold $\Leftarrow$ 3. 56% $\langle$

#### Example 5 FUSION PROTEIN PREPARATION

##### 10 a. GPCR:Gs Fusion Construct

The design of the constitutively activated GPCR-G protein fusion construct was accomplished as follows: both the 5' and 3' ends of the rat G protein Gs $\alpha$  (long form; Itoh, H. et al., 83 *PNAS* 3776 (1986)) were engineered to include a HindIII (5'-  
15 AAGCTT-3') sequence thereon. Following confirmation of the correct sequence (including the flanking HindIII sequences), the entire sequence was shuttled into pcDNA3.1(-) (Invitrogen, cat. no. V795-20) by subcloning using the HindIII restriction site of that vector. The correct orientation for the Gs $\alpha$  sequence was determined after

subcloning into pcDNA3.1(-). The modified pcDNA3.1(-) containing the rat Gs $\alpha$  gene at HindIII sequence was then verified; this vector was now available as a "universal" Gs $\alpha$  protein vector. The pcDNA3.1(-) vector contains a variety of well-known restriction sites upstream of the HindIII site, thus beneficially providing the ability to  
5 insert, upstream of the Gs protein, the coding sequence of an endogenous, constitutively active GPCR. This same approach can be utilized to create other "universal" G protein vectors, and, of course, other commercially available or proprietary vectors known to the artisan can be utilized – the important criteria is that the sequence for the GPCR be upstream and in-frame with that of the G protein.

10 RUP13 couples via Gs. For the following exemplary GPCR Fusion Proteins, fusion to Gs $\alpha$  was accomplished.

A RUP13-Gs $\alpha$  Fusion Protein construct was made as follows: primers were designed as follows:

5'-gatc[TCTAGAAT]GGAGTCCTACCCATCCCCCAG -3' (SEQ.ID.NO.:97; sense)

15 5'-gatc[GATATC]CGTGACTCCAGCCGGGGTGAGGCGGC-3' (SEQ.ID.NO.:98; antisense).

Nucleotides in lower caps are included as spacers in the restriction sites (designated in brackets) between the G protein and RUP13. The sense and anti-sense primers included the restriction sites for XbaI and EcoRV, respectively, such that spacers (attributed to the restriction sites) exists between the G protein and RUP15.

20 PCR was then utilized to secure the respective receptor sequences for fusion within the Gs $\alpha$  universal vector disclosed above, using the following protocol for each: 100ng cDNA for RUP15 was added to separate tubes containing 2 $\mu$ L of each primer (sense and anti-sense), 3 $\mu$ L of 10mM dNTPs, 10 $\mu$ L of 10XTaqPlus™ Precision buffer, 1 $\mu$ L of TaqPlus™ Precision polymerase (Stratagene: #600211), and 80 $\mu$ L of water.  
25 Reaction temperatures and cycle times for RUP15 were as follows with cycle steps 2



through 4 were repeated 35 times: 94°C for 1 min; 94°C for 30 seconds; 62°C for 20 sec; 72°C 1 min 40sec; and 72° C 5 min . PCR product for was run on a 1% agarose gel and then purified (data not shown). The purified product was digested with XbaI and EcoRV and the desired inserts purified and ligated into the Gs universal vector at the  
5 respective restriction site. The positive clones was isolated following transformation and determined by restriction enzyme digest; expression using 293 cells was accomplished following the protocol set forth *infra*. Each positive clone for RUP15-Gs Fusion Protein was sequenced to verify correctness. (See, SEQ.ID.NO.:99 for nucleic acid sequence and SEQ.ID.NO.:100 for amino acid sequence ).

10 RUP15 couples via Gs. For the following exemplary GPCR Fusion Proteins, fusion to Gs $\alpha$  was accomplished.

A RUP15-Gs $\alpha$  Fusion Protein construct was made as follows: primers were designed as follows:

5'-TCTAGAATGACGTCCACCTGCACCAACAGC-3' (SEQ.ID.NO.:101; sense)

15 5'-gatatcGCAGGAAAAGTAGCAGAATCGTAGGAAG-3' (SEQ.ID.NO.:102; antisense).

Nucleotides in lower caps are included as spacers in the restriction sites between the G protein and RUP15. The sense and anti-sense primers included the restriction sites for EcoRV and XbaI, respectively, such that spacers (attributed to the restriction sites) exists between the G protein and RUP15.

20 PCR was then utilized to secure the respective receptor sequences for fusion within the Gs $\alpha$  universal vector disclosed above, using the following protocol for each: 100ng cDNA for RUP15 was added to separate tubes containing 2 $\mu$ l of each primer (sense and anti-sense), 3 $\mu$ L of 10mM dNTPs, 10 $\mu$ L of 10XTaqPlus™ Precision buffer, 1 $\mu$ L of TaqPlus™ Precision polymerase (Stratagene: #600211), and 80 $\mu$ L of water.  
25 Reaction temperatures and cycle times for RUP15 were as follows with cycle steps 2

through 4 were repeated 35 times: 94°C for 1 min; 94°C for 30 seconds; 62°C for 20 sec; 72°C 1 min 40sec; and 72° C 5 min . PCR product for was run on a 1% agarose gel and then purified (data not shown). The purified product was digested ). The purified product was digested with EcoRV and Xba1 and the desired inserts purified and

5 ligated into the Gs universal vector at the respective restriction site. The positive clones was isolated following transformation and determined by restriction enzyme digest; expression using 293 cells was accomplished following the protocol set forth *infra*. Each positive clone for RUP15-Gs Fusion Protein was sequenced to verify correctness. (See, SEQ.ID.NO.:103 for nucleic acid sequence and SEQ.ID.NO.:104 for amino acid

10 sequence ).

**b. Gq(6 amino acid deletion)/Gi Fusion Construct**

The design of a Gq (del)/Gi fusion construct can be accomplished as follows: the N-terminal six (6) amino acids (amino acids 2 through 7, having the sequence of TLESIM (SEQ.ID.NO.: 129) Gαq-subunit will be deleted and the C-terminal five (5)

15 amino acids, having the sequence EYNLV (SEQ.ID.NO.:130) will be replace with the corresponding amino acids of the Gαi Protein, having the sequence DCGLF (SEQ.ID.NO.:131). This fusion construct will be obtained by PCR using the following primers:

20 5'-gatcaagcttcCATGGCGTGCTGCCTGAGCGAGGAG-3' (SEQ.ID.NO.:132) and

5'-gatcggatccTTAGAACAGGCCGCAGTCCTTCAGGTTTCAGCTGCAGGATGGTG-3' (SEQ.ID.NO.:133)

and Plasmid 63313 which contains the mouse Gαq-wild type version with a

25 hemagglutinin tag as template. Nucleotides in lower caps are included as spacers.

TaqPlus Precision DNA polymerase (Stratagene) will be utilized for the amplification by the following cycles, with steps 2 through 4 repeated 35 times: 95°C

for 2 min; 95°C for 20 sec; 56°C for 20 sec; 72°C for 2 min; and 72°C for 7 min. The PCR product will be cloned into a pCRII-TOPO vector (Invitrogen) and sequenced using the ABI Big Dye Terminator kit (P.E. Biosystem). Inserts from a TOPO clone containing the sequence of the fusion construct will be shuttled into the expression  
 5 vector pcDNA3.1(+) at the HindIII/BamHI site by a 2 step cloning process.

### Example 6

#### TISSUE DISTRIBUTION OF THE DISCLOSED HUMAN GPCRS: RT-PCR

RT-PCR was applied to confirm the expression and to determine the tissue  
 10 distribution of several novel human GPCRs. Oligonucleotides utilized were GPCR-specific and the human multiple tissue cDNA panels (MTC, Clontech) as templates. Taq DNA polymerase (Stratagene) were utilized for the amplification in a 40µl reaction according to the manufacturer's instructions. 20µl of the reaction will be loaded on a 1.5% agarose gel to analyze the RT-PCR products. Table J below lists the  
 15 receptors, the cycle conditions and the primers utilized.

**TABLE J**

Receptor Identifier	Cycle Conditions Min ('), Sec (") Cycles 2-4 repeated 30 times	5' Primer (SEQ.ID.NO.)	3' Primer (SEQ.ID.NO.)	DNA Fragment	Tissue Expression
hRUP10	94° for 30" 94° for 10" 62°C for 20" 72° for 1' 72° for 7' *cycles 2-4 repeated 35 times	CATGTATGC CAGCGTCCT GCTCC (105)	GCTATGCCTG AAGCCAGTC TTGTG (106)	730bp	Kidney, leukocyte, liver, placenta and spleen
hRUP11	94° for 2' 94° for 15" 67°C for 15" 72° for 45" 72° f r 5'	GCACCTGCT CCTGAGCAC CTTCTCC (107)	CACAGCGCT GCAGCCCTG CAGCTGGC (108)	630bp	Liver, kidney, pancreas, colon, small intestinal, spleen and prostate

<b>hRUP12</b>	94° for 2' 94° for 15" 66°C for 15" 72° for 45" 72° for 5'	CCAGTGATG ACTCTGTCC AGCCTG (109)	CAGACACTT GGCAGGGAC GAGGTG (110)	490bp	Brain, c lon, heart, kidney, leukocyte, pancreas, prostate, small intestinal, spleen, testis, and thymus
---------------	--	--	--	-------	---

<b>hRUP13</b>	94° for 1' 94° for 15" 68°C for 20" 72° for 1' 45" 72° for 5'	CTTGTGGTCT ACTGCAGCA TGTTCCG (111)	CATATCCCTC CGAGTGTCC AGCGGC (112)	700bp	Placenta and lung
<b>hRUP14</b>	94° for 1' 94° for 15" 68°C for 20" 72° for 1' 45" 72° for 5'	ATGGATCCT TATCATGGC TTCCTC (113)	CAAGAACAG GTCTCATCTA AGAGCTCC (114)	700bp	Not yet determined
<b>hRUP16</b>	94° for 30" 94° for 5" 69°C for 15" 72° for 30" 72° for 5'	CTCTGATGC CATCTGCTG GATTCCTG (115)	GTAGTCCACT GAAAGTCCA GTGATCC (116)	370bp	Fetal brain, fetal kidney and fetal skeletal muscle
<b>hRUP18</b>	94° for 2' 94° for 15" 60°C for 20" 72° for 1' 72° for 5'	TGGTGGCGA TGGCCAACA GCGCTC (117)	GTTGCGCCTT AGCGACAGA TGACC (118)	330bp	Pancreas
<b>hRUP21</b>	94° for 1' 94° for 15" 56°C for 20" 72° for 40" *cycles 2-3 repeated 30 times	TCAACCTGT ATAGCAGCA TCCTC (119)	AAGGAGTAG CAGAATGGT TAGCC (120)		Kidney, lung and testis
<b>hRUP22</b>	94° for 30" 94° for 15" 69°C for 20" 72° for 40" *cycles 2-3 repeated 30 times	GACACCTGT CAGCGGTCG TGTGTG (121)	CTGATGGAA GTAGAGGCT GTCCATCTC (122)		Testis, thymus and spleen
<b>hRUP23</b>	94° for 2' 94° for 15" 60°C for 20" 72° for 1' 72° for 5'	GCGCTGAGC GCAGACCAG TGGCTG (123)	CACGGTGAC GAAGGGCAC GAGCTC (124)	520bp	Placenta
<b>hRUP26</b>	94° for 2' 94° for 15" 65°C for 20" 72° for 1' 72° for 5'	AGCCATCCC TGCCAGGAA GCATGG (125)	CCAGGTAGG TGTGCAGCA CAATGGC (126)	470bp	Pancreas
<b>hRUP27</b>	94° for 30" 94° for 10" 55°C for 20" 72° for 1' 72° for 3' *cycles 2-4 repeated 35 times	CTGTTCAAC AGGGCTGGT TGGCAAC (127)	ATCATGTCTA GACTCATGGT GATCC (128)	890bp	Brain

**Example 7****Protocol: Direct Identification of Inverse Agonists and Agonists****A. [<sup>35</sup>S]GTPγS Assay**

5           Although we have utilized endogenous, constitutively active GPCRs for the direct identification of candidate compounds as, *e.g.*, inverse agonists, for reasons that are not altogether understood, intra-assay variation can become exacerbated. Preferably, then, a GPCR Fusion Protein, as disclosed above, is also utilized with a non-endogenous, constitutively activated GPCR. We have determined that when such a  
10   protein is used, intra-assay variation appears to be substantially stabilized, whereby an effective signal-to-noise ratio is obtained. This has the beneficial result of allowing for a more robust identification of candidate compounds. Thus, it is preferred that for direct identification, a GPCR Fusion Protein be used and that when utilized, the following assay protocols be utilized.

**15                   1.     Membrane Preparation**

Membranes comprising the constitutively active orphan GPCR Fusion Protein of interest and for use in the direct identification of candidate compounds as inverse agonists, agonists or partial agonists are preferably prepared as follows:

**a.     Materials**

20           “Membrane Scrape Buffer” is comprised of 20mM HEPES and 10mM EDTA, pH 7.4; “Membrane Wash Buffer” is comprised of 20 mM HEPES and 0.1 mM EDTA, pH 7.4; “Binding Buffer” is comprised of 20mM HEPES, 100 mM NaCl, and 10 mM MgCl<sub>2</sub>, pH 7.4

**b.     Procedure**

25           All materials will be kept on ice throughout the procedure. Firstly, the media will be aspirated from a confluent monolayer of cells, followed by rinse with 10ml cold

PBS, followed by aspiration. Thereafter, 5ml of Membrane Scrape Buffer will be added to scrape cells; this will be followed by transfer of cellular extract into 50ml centrifuge tubes (centrifuged at 20,000 rpm for 17 minutes at 4°C). Thereafter, the supernatant will be aspirated and the pellet will be resuspended in 30ml Membrane Wash Buffer  
5 followed by centrifuge at 20,000 rpm for 17 minutes at 4°C. The supernatant will then be aspirated and the pellet resuspended in Binding Buffer. This will then be homogenized using a Brinkman polytron™ homogenizer (15-20 second bursts until the all material is in suspension). This is referred to herein as "Membrane Protein".

## **2. Bradford Protein Assay**

10 Following the homogenization, protein concentration of the membranes will be determined using the Bradford Protein Assay (protein can be diluted to about 1.5mg/ml, aliquoted and frozen (-80°C) for later use; when frozen, protocol for use will be as follows: on the day of the assay, frozen Membrane Protein is thawed at room temperature, followed by vortex and then homogenized with a polytron at about  
15 12 x 1,000 rpm for about 5-10 seconds; it was noted that for multiple preparations, the homogenizer should be thoroughly cleaned between homogenization of different preparations).

### **a. Materials**

Binding Buffer (as per above); Bradford Dye Reagent; Bradford Protein  
20 Standard will be utilized, following manufacturer instructions (Biorad, cat. no. 500-0006).

### **b. Procedure**

Duplicate tubes will be prepared, one including the membrane, and one as a control "blank". Each contained 800ul Binding Buffer. Thereafter, 10µl of Bradford  
25 Protein Standard (1mg/ml) will be added to each tube, and 10µl of membrane Protein

will then be added to just one tube (not the blank). Thereafter, 200ul of Bradford Dye Reagent will be added to each tube, followed by vortex of each. After five (5) minutes, the tubes will be re-vortexed and the material therein will be transferred to cuvettes. The cuvettes will then be read using a CECIL 3041 spectrophotometer, at  
5 wavelength 595.

### 3. Direct Identification Assay

#### a. Materials

GDP Buffer consisted of 37.5 ml Binding Buffer and 2mg GDP (Sigma, cat. no. G-7127), followed by a series of dilutions in Binding Buffer to obtain 0.2  $\mu$ M GDP  
10 (final concentration of GDP in each well was 0.1  $\mu$ M GDP); each well comprising a candidate compound, has a final volume of 200ul consisting of 100 $\mu$ l GDP Buffer (final concentration, 0.1 $\mu$ M GDP), 50ul Membrane Protein in Binding Buffer, and 50 $\mu$ l [ $^{35}$ S]GTP $\gamma$ S (0.6 nM) in Binding Buffer (2.5  $\mu$ l [ $^{35}$ S]GTP $\gamma$ S per 10ml Binding Buffer).

#### b. Procedure

15 Candidate compounds will be preferably screened using a 96-well plate format (these can be frozen at -80°C). Membrane Protein (or membranes with expression vector excluding the GPCR Fusion Protein, as control), will be homogenized briefly until in suspension. Protein concentration will then be determined using the Bradford Protein Assay set forth above. Membrane Protein (and control) will then be diluted to  
20 0.25mg/ml in Binding Buffer (final assay concentration, 12.5 $\mu$ g/well). Thereafter, 100  $\mu$ l GDP Buffer was added to each well of a Wallac Scintistrip™ (Wallac). A 5ul pin-tool will then be used to transfer 5  $\mu$ l of a candidate compound into such well (*i.e.*, 5 $\mu$ l in total assay volume of 200  $\mu$ l is a 1:40 ratio such that the final screening concentration of the candidate compound is 10 $\mu$ M). Again, to avoid contamination, after each transfer  
25 step the pin tool should be rinsed in three reservoirs comprising water (1X), ethanol (1X)



and water (2X) – excess liquid should be shaken from the tool after each rinse and dried with paper and kimwipes. Thereafter, 50 µl of Membrane Protein will be added to each well (a control well comprising membranes without the GPCR Fusion Protein was also utilized), and pre-incubated for 5-10 minutes at room temperature. Thereafter, 50µl of  
5 [35S]GTPγS (0.6 nM) in Binding Buffer will be added to each well, followed by incubation on a shaker for 60 minutes at room temperature (again, in this example, plates were covered with foil). The assay will then be stopped by spinning of the plates at 4000 RPM for 15 minutes at 22°C. The plates will then be aspirated with an 8 channel manifold and sealed with plate covers. The plates will then be read on a Wallacc 1450  
10 using setting “Prot. #37” (as per manufacturer instructions).

#### B. Cyclic AMP Assay

Another assay approach to directly identified candidate compound was accomplished by utilizing a cyclase-based assay. In addition to direct identification, this  
15 assay approach can be utilized as an independent approach to provide confirmation of the results from the [35S]GTPγS approach as set forth above.

A modified Flash Plate™ Adenylyl Cyclase kit (New England Nuclear, Cat. No. SMP004A) was preferably utilized for direct identification of candidate compounds as inverse agonists and agonists to constitutively activated orphan GPCRs in accordance  
20 with the following protocol.

Transfected cells were harvested approximately three days after transfection. Membranes were prepared by homogenization of suspended cells in buffer containing 20mM HEPES, pH 7.4 and 10mM MgCl<sub>2</sub>. Homogenization was performed on ice using a Brinkman Polytron™ for approximately 10 seconds. The resulting homogenate is  
25 centrifuged at 49,000 X g for 15 minutes at 4°C. The resulting pellet was then resuspended in buffer containing 20mM HEPES, pH 7.4 and 0.1 mM EDTA,

homogenized for 10 seconds, followed by centrifugation at 49,000 X g for 15 minutes at 4°C. The resulting pellet was then stored at -80°C until utilized. On the day of direct identification screening, the membrane pellet was slowly thawed at room temperature, resuspended in buffer containing 20mM HEPES, pH 7.4 and 10mM MgCL<sub>2</sub>, to yield a  
5 final protein concentration of 0.60mg/ml (the resuspended membranes are placed on ice until use).

cAMP standards and Detection Buffer (comprising 2 µCi of tracer [<sup>125</sup>I] cAMP (100 µl] to 11 ml Detection Buffer) were prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer was prepared fresh for screening and  
10 contained 20mM HEPES, pH 7.4, 10mM MgCl<sub>2</sub>, 20mM phosphocreatine (Sigma), 0.1 units/ml creatine phosphokinase (Sigma), 50 µM GTP (Sigma), and 0.2 mM ATP (Sigma); Assay Buffer was then stored on ice until utilized.

Candidate compounds identified as per above (if frozen, thawed at room temperature) were added, preferably, to 96-well plate wells (3µl/well; 12µM final assay  
15 concentration), together with 40 µl Membrane Protein (30µg/well) and 50µl of Assay Buffer. This admixture was then incubated for 30 minutes at room temperature, with gentle shaking.

Following the incubation, 100µl of Detection Buffer was added to each well, followed by incubation for 2-24 hours. Plates were then counted in a Wallac  
20 MicroBeta™ plate reader using "Prot. #31" (as per manufacturer instructions).

A representative screening assay plate (96 well format) result is presented in Figure 12. Each bar represents the results for a different compound in each well, plus RUP13-Gsα Fusion Protein construct, as prepared in Example 5(a) above. The representative results presented in Figure 12 also provide standard deviations based upon  
25 the mean results of each plate ("m") and the mean plus two arbitrary preference for

selection of inverse agonists as "leads" from the primary screen involves selection of candidate compounds that that reduce the per cent response by at least the mean plate response, minus two standard deviations. Conversely, an arbitrary preference for selection of an agonists as "leads" from the primary screen involves selection of

5 candidate compounds that increase the per cent response by at least the mean plate response, plus the two standard deviations. Based upon these selection processes, the candidate compounds in the following wells were directly identified as putative inverse agonist (Compound A) and agonist (Compound B) to RUP13 in wells A2 and G9, respectively. *See*, Figure 12. It is noted for clarity: these compounds have been directly

10 identified without any knowledge of the endogenous ligand for this GPCR. By focusing on assay techniques that are based upon receptor function, and not compound binding affinity, we are able to ascertain compounds that are able to reduce the functional activity of this receptor (Compound A) as well as increase the functional activity of the receptor (Compound B). Based upon the location of these receptor in lung tissue (*see*,

15 for example, hRUP13 and hRUP21 in Example 6), pharmaceutical agents can be developed for potential therapeutic treatment of lung cancer.

References cited throughout this patent document, including co-pending and related patent applications, unless otherwise indicated, are fully incorporated herein by reference. Modifications and extension of the disclosed inventions that are within the

20 purview of the skilled artisan are encompassed within the above disclosure and the claims that follow.

Although a variety of expression vectors are available to those in the art, for purposes of utilization for both the endogenous and non-endogenous human GPCRs, it is most preferred that the vector utilized be pCMV. This vector was deposited with the

25 American Type Culture Collection (ATCC) on October 13, 1998 (10801 University

Blvd., Manassas, VA 20110-2209 USA) under the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure. The DNA was tested by the ATCC and determined to be viable. The ATCC has assigned the following deposit number to pCMV: ATCC #203351.

5 //

//

//

//

//

10 //

//

//

//

//

15 //

//

//

//

//

20 //

//

//

//

//

25 //

## CLAIMS

What is claimed is:

1. A G protein-coupled receptor encoded by an amino acid sequence of  
5 SEQ.ID.NO.:2.
2. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 1.
3. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:1.
4. A host cell comprising the plasmid of claim 3.
- 10 5. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:4.
6. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 5.
7. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:3.
- 15 8. A host cell comprising the plasmid of claim 7.
9. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:6.
10. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 9.
- 20 11. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:5.
12. A host cell comprising the plasmid of claim 11.
13. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:8.
14. A non-endogenous, constitutively activated version of the G protein-coupled  
25 receptor of claim 13.

15. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:7.
16. A host cell comprising the plasmid of claim 15.
17. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:10.
- 5 18. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 17.
19. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:9.
20. A host cell comprising the plasmid of claim 19.
21. A G protein-coupled receptor encoded by an amino acid sequence of  
10 SEQ.ID.NO.:12.
22. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 21 comprising an amino acid sequence of SEQ.ID.NO.84.
23. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:11.
24. A host cell comprising the plasmid of claim 23.
- 15 25. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:14.
26. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 25 comprising an amino acid sequence of SEQ.ID.NO.88.
27. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:13.
- 20 28. A host cell comprising the plasmid of claim 27.
29. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:16.
30. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor f claim 29 comprising an amino acid sequence of SEQ.ID.NO.:92.
- 25 31. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:15.

32. A host cell comprising the plasmid of claim 31.
33. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:18.
34. A non-endogenous, constitutively activated version of the G protein-coupled  
5 receptor of claim 33.
35. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:17.
36. A host cell comprising the plasmid of claim 35.
37. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:20.
- 10 38. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 37.
39. A plasmid comprising a vector and the cDNA of SE.ID.NO.:19.
40. A host cell comprising the plasmid of claim 39.
41. A G protein-coupled receptor encoded by an amino acid sequence of  
15 SEQ.ID.NO.:22.
42. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 41.
43. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:21.
44. A host cell comprising the plasmid of claim 43.
- 20 45. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:24.
46. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 45.
47. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:23.
- 25 48. A host cell comprising the plasmid of claim 47.

49. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:26.
50. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 49.
- 5 51. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:25.
52. A host cell comprising the plasmid of claim 51.
53. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:28.
54. A non-endogenous, constitutively activated version of the G protein-coupled  
10 receptor of claim 53.
55. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:27.
56. A host cell comprising the plasmid of claim 55.
57. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:30.
- 15 58. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 57.
59. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:29.
60. A host cell comprising the plasmid of claim 59.
61. A G protein-coupled receptor encoded by an amino acid sequence of  
20 SEQ.ID.NO.:32.
62. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 61 comprising an amino acid sequence of SEQ.ID.NO.:96.
63. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:95.
64. A host cell comprising the plasmid of claim 63.



65. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:34.
66. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 65.
- 5 67. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:33.
68. A host cell comprising the plasmid of claim 67.
69. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:36.
70. A non-endogenous, constitutively activated version of the G protein-coupled  
10 receptor of claim 69.
71. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:35.
72. A host cell comprising the plasmid of claim 71.
73. A G protein-coupled receptor encoded by an amino acid sequence of  
SEQ.ID.NO.:38.
- 15 74. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 73.
75. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:37.
76. A host cell comprising the plasmid of claim 75.
77. A G protein-coupled receptor encoded by an amino acid sequence of  
20 SEQ.ID.NO.:40.
78. A non-endogenous, constitutively activated version of the G protein-coupled  
receptor of claim 77.
79. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:39.
80. A host cell comprising the plasmid of claim 79.

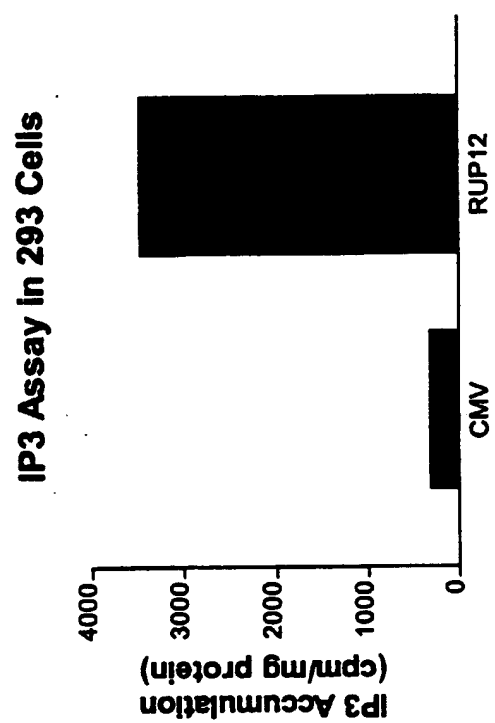


Figure 1

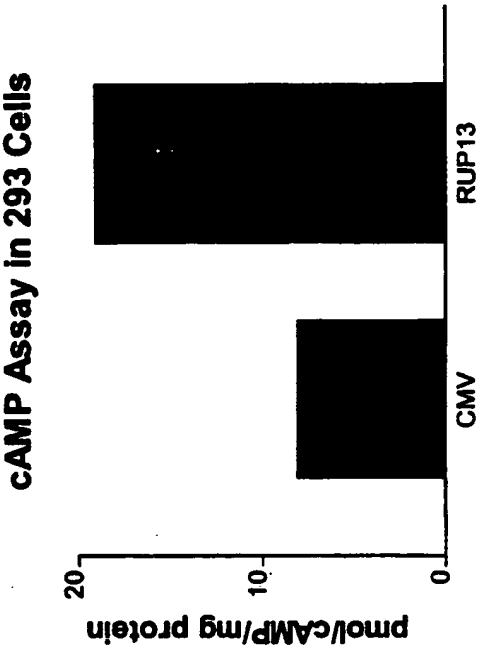


Figure 2

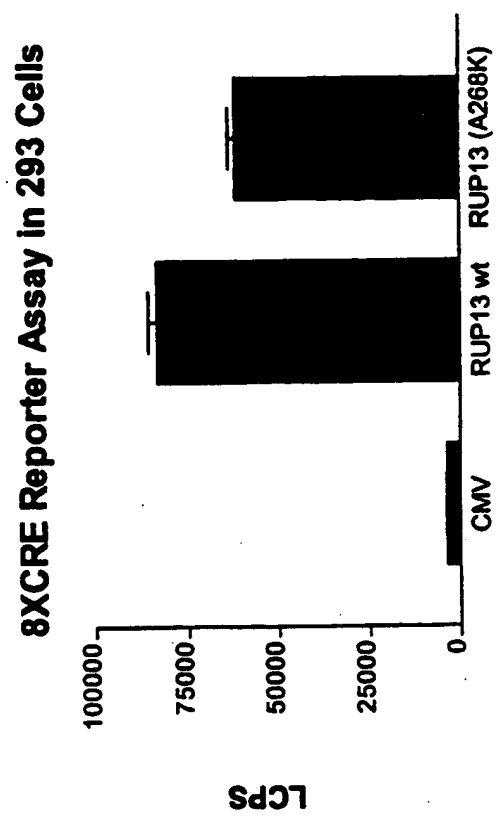


Figure 3

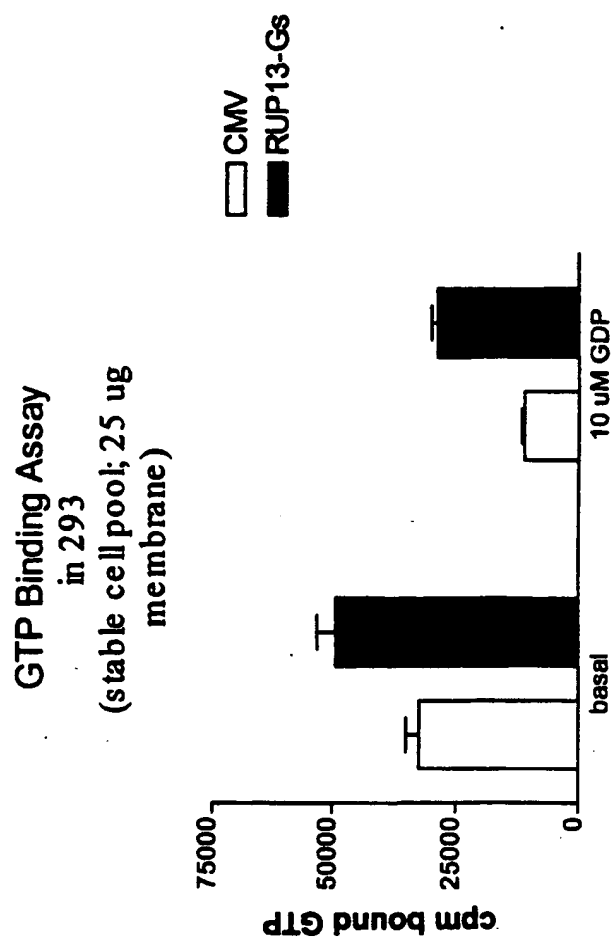


Figure 4

8XCRE Reporter Assay in 293 Cells

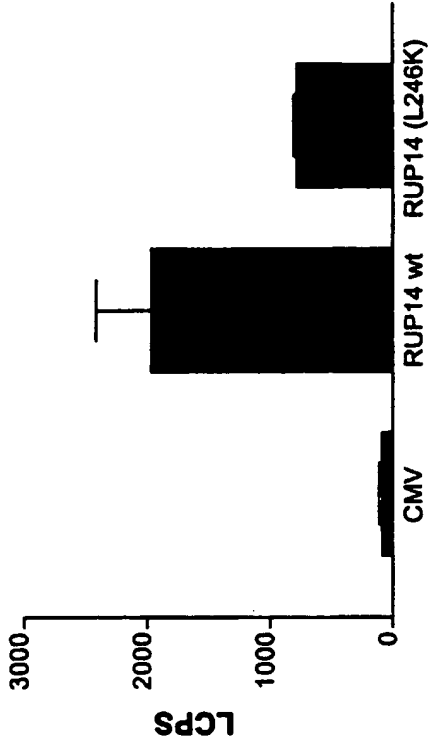


Figure 5

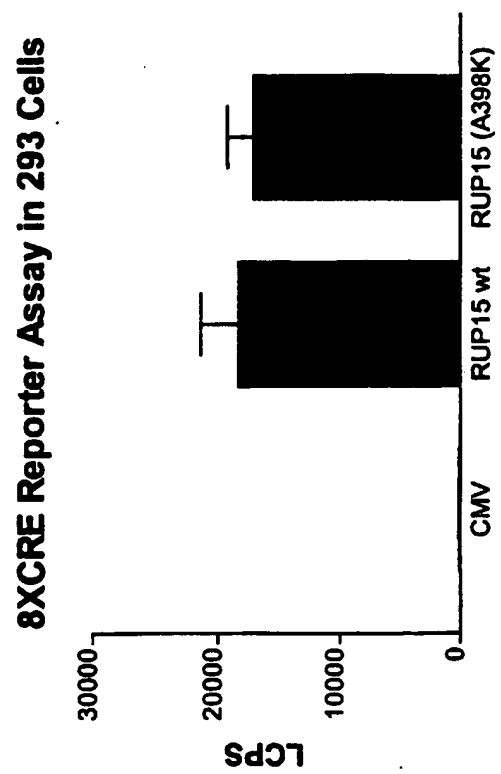


Figure 6

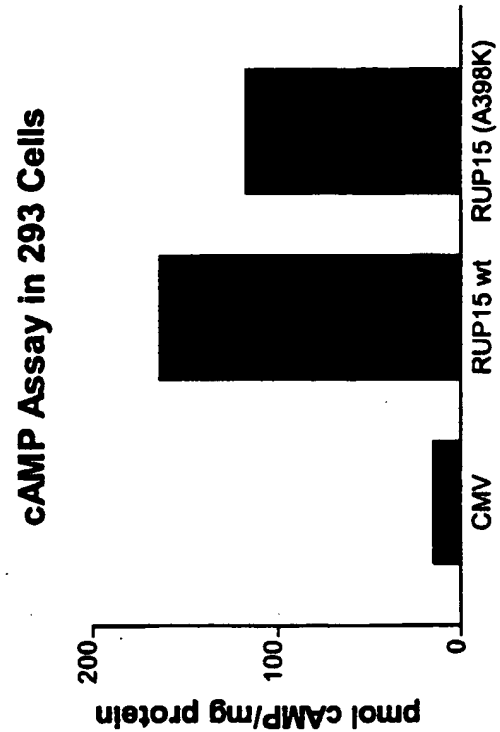


Figure 7



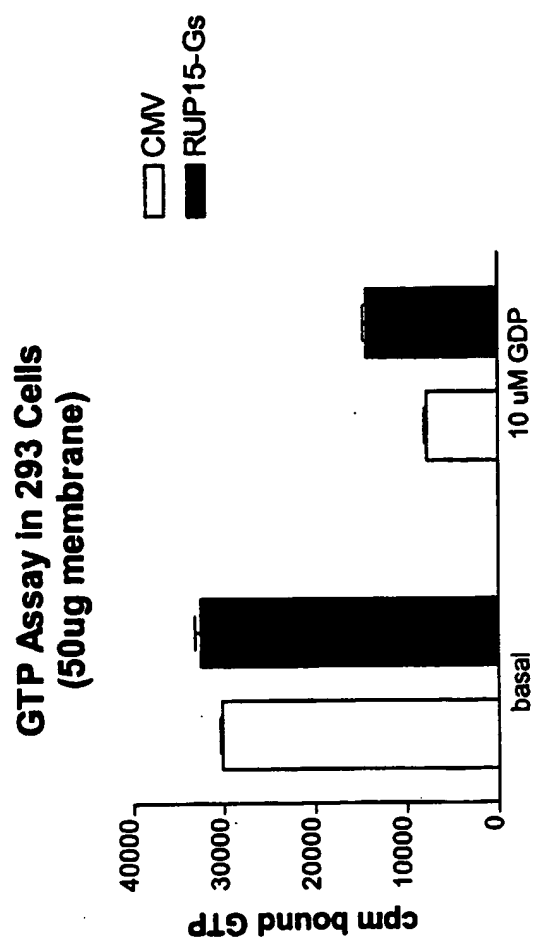


Figure 8

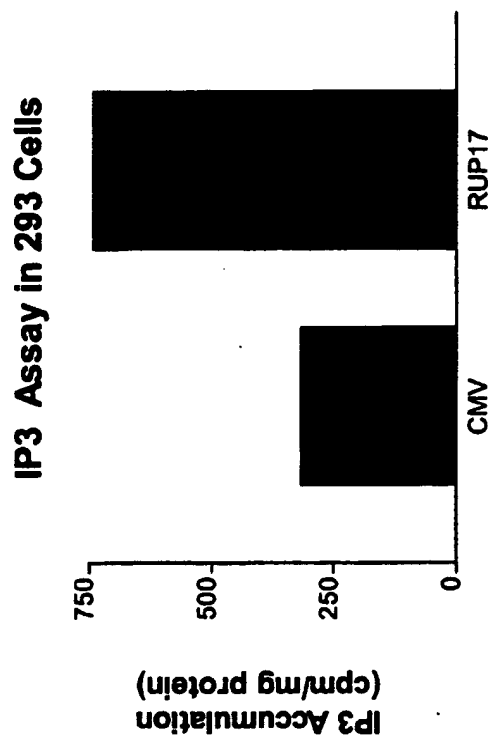


Figure 9

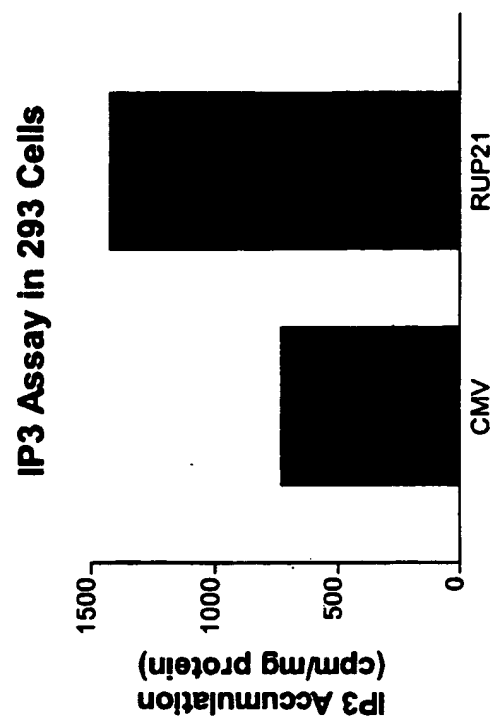


Figure 10

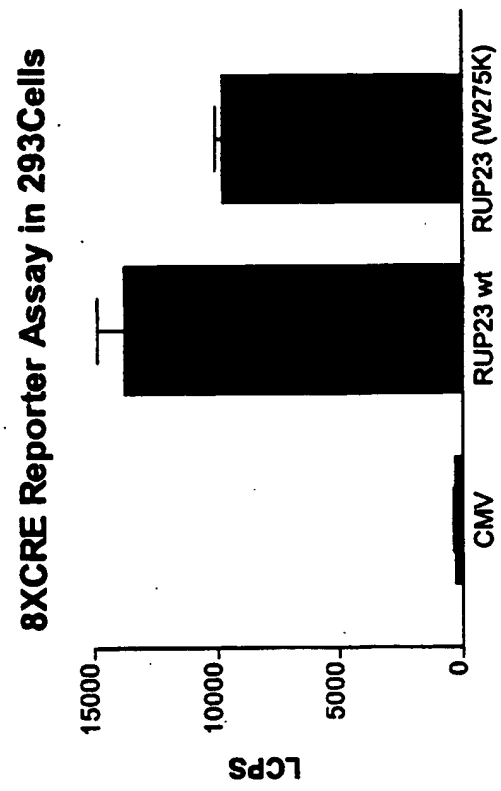


Figure 11

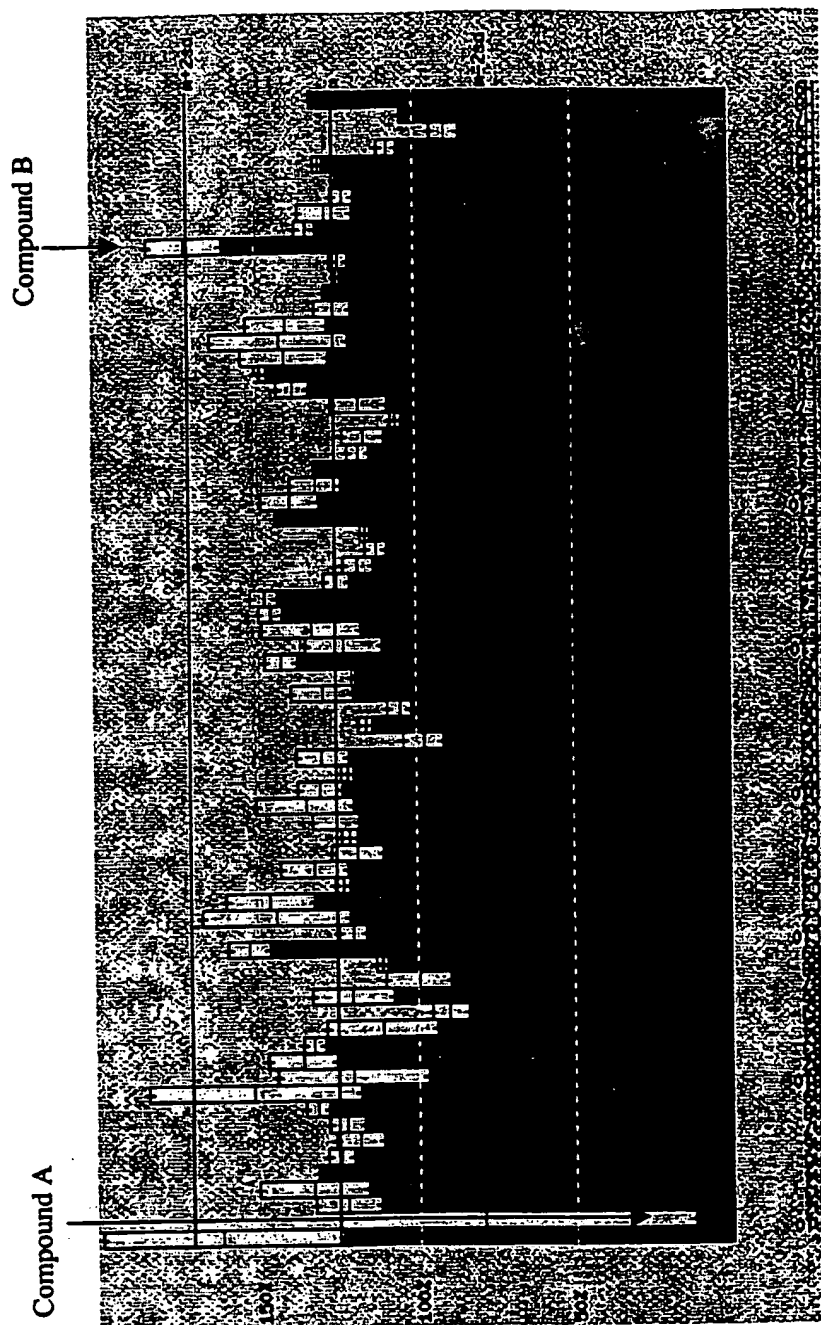


Figure 12

## SEQUENCE LISTING

<110> Arena Pharmaceuticals, Inc.  
Chen, Rupong  
Dang, Huong T.  
Lowitz, Kevin P.

<120> Non-Endogenous, Constitutively Activated Human G Protein-Coupled Receptors

<130> AREN0087

<150> 60/166,088  
<151> 1999-11-17

<150> 60/166,369  
<151> 1999-11-17

<150> 60/166,099  
<151> 1999-11-17

<150> 61/171,902  
<151> 1999-12-23

<150> 60/171,901  
<151> 1999-12-23

<150> 60/171,900  
<151> 1999-12-23

<150> 60/181,749  
<151> 2000-02-11

<150> 60/189,258  
<151> 2000-03-14

<150> 60/189,259  
<151> 2000-03-14

<150> 60/195,899  
<151> 2000-04-10

<150> 60/196,078  
<151> 2000-04-10

<150> 60/195,898  
<151> 2000-04-10

<150> 60/200,419  
<151> 2000-04-28

<150> 60/203,630  
<151> 2000-05-12

<150> 60/210,741  
<151> 2000-06-12

<150> 60/210,982  
<151> 2000-06-12

<150> 60/226,760  
<151> 2000-08-21

<150> 60/235,779  
<151> 2000-09-26

<150> 60/235,418  
<151> 2000-09-26

<150> 60/242,332  
<151> 2000-10-20

<150> 60/242,343  
<151> 2000-10-20

<150> 60/243,019  
<151> 2000-10-24

<160> 133

<170> PatentIn version 3.0

<210> 1  
<211> 1155  
<212> DNA  
<213> Homo sapiens

```

<400> 1
atggcagccc agaatggaaa caccagtttc acacccaact ttaatccacc ccaagaccat    60
gcctcctccc tctcctttaa cttcagttat ggtgattatg acctccctat ggatgaggat    120
gaggacatga ccaagaccgg gaccttcttc gcagccaaga tcgtcattgg cattgcactg    180
gcaggcatca tgctggctcg cggcatcggt aactttgtct ttatcgctgc cctcaccggc    240
tataagaagt tgcgcaacct caccaatctg ctcatggcca acctggccat ctccgacttc    300
ctggtggcca tcatctgctg ccccttcgag atggactact acgtggtagc gcagctctcc    360
tgggagcatg gccacgtgct ctgtgcctcc gtcaactacc tgcgcaccgt ctccctctac    420
gtctccacca atgccttgct ggccattgcc attgacagat atctcgccat cgttcacccc    480
ttgaaaccac ggatgaatta tcaaacggcc tccttcctga tcgccttggc ctggatggtg    540
tccattctca ttgccatccc atcgggttac ttgcaacag aaacggctct ctttattgtc    600
aagagccagg agaagatctt ctgtggccag atctggcctg tggatcagca gctctactac    660
aagtcctact tcctcttcat ctttgggtgc gagttcgtgg gccctgtggt caccatgacc    720
ctgtgctatg ccaggatctc ccgggagctc tggttcaagg cagtccctgg gttccagacg    780
gagcagattc gcaagcggct gcgctgccgc aggaagacgg tcctggtgct catgtgcatt    840
ctcacggcct atgtgctgtg ctgggcaccc ttctacggtt tcaccatcgt tcgtgacttc    900
ttccccactg tgttcgtgaa ggaaaagcac tacctcactg ccttctacgt ggtcgagtgc    960
atcgccatga gcaacagcat gatcaacacc gtgtgcttcg tgacggtcaa gaacaacacc   1020
atgaagtact tcaagaagat gatgctgctg cactggcgct cctcccagcg ggggagcaag   1080
tccagtgtcg accttgacct cagaaccaac ggggtgccca ccacagaaga ggtggactgt   1140
atcaggctga agtga                                     1155

```

<210> 2  
<211> 384

<212> PRT  
 <213> Homo sapiens

<400> 2

```

Met Ala Ala Gln Asn Gly Asn Thr Ser Phe Thr Pro Asn Phe Asn Pro
1          5          10          15
Pro Gln Asp His Ala Ser Ser Leu Ser Phe Asn Phe Ser Tyr Gly Asp
          20          25          30
Tyr Asp Leu Pro Met Asp Glu Asp Glu Asp Met Thr Lys Thr Arg Thr
          35          40          45
Phe Phe Ala Ala Lys Ile Val Ile Gly Ile Ala Leu Ala Gly Ile Met
50          55          60
Leu Val Cys Gly Ile Gly Asn Phe Val Phe Ile Ala Ala Leu Thr Arg
65          70          75          80
Tyr Lys Lys Leu Arg Asn Leu Thr Asn Leu Leu Ile Ala Asn Leu Ala
          85          90          95
Ile Ser Asp Phe Leu Val Ala Ile Ile Cys Cys Pro Phe Glu Met Asp
          100          105          110
Tyr Tyr Val Val Arg Gln Leu Ser Trp Glu His Gly His Val Leu Cys
          115          120          125
Ala Ser Val Asn Tyr Leu Arg Thr Val Ser Leu Tyr Val Ser Thr Asn
          130          135          140
Ala Leu Leu Ala Ile Ala Ile Asp Arg Tyr Leu Ala Ile Val His Pro
145          150          155          160
Leu Lys Pro Arg Met Asn Tyr Gln Thr Ala Ser Phe Leu Ile Ala Leu
          165          170          175
Val Trp Met Val Ser Ile Leu Ile Ala Ile Pro Ser Ala Tyr Phe Ala
          180          185          190
Thr Glu Thr Val Leu Phe Ile Val Lys Ser Gln Glu Lys Ile Phe Cys
          195          200          205
Gly Gln Ile Trp Pro Val Asp Gln Gln Leu Tyr Tyr Lys Ser Tyr Phe
          210          215          220
Leu Phe Ile Phe Gly Val Glu Phe Val Gly Pro Val Val Thr Met Thr
225          230          235          240
Leu Cys Tyr Ala Arg Ile Ser Arg Glu Leu Trp Phe Lys Ala Val Pro
          245          250          255
Gly Phe Gln Thr Glu Gln Ile Arg Lys Arg Leu Arg Cys Arg Arg Lys
          260          265          270
Thr Val Leu Val Leu Met Cys Ile Leu Thr Ala Tyr Val Leu Cys Trp
          275          280          285
Ala Pro Phe Tyr Gly Phe Thr Ile Val Arg Asp Phe Phe Pro Thr Val
          290          295          300
Phe Val Lys Glu Lys His Tyr Leu Thr Ala Phe Tyr Val Val Glu Cys
305          310          315          320

```



Ile Ala Met Ser Asn Ser Met Ile Asn Thr Val Cys Phe Val Thr Val  
325 330 335

Lys Asn Asn Thr Met Lys Tyr Phe Lys Lys Met Met Leu Leu His Trp  
340 345 350

Arg Pro Ser Gln Arg Gly Ser Lys Ser Ser Ala Asp Leu Asp Leu Arg  
355 360 365

Thr Asn Gly Val Pro Thr Thr Glu Glu Val Asp Cys Ile Arg Leu Lys  
370 375 380

<210> 3  
<211> 1260  
<212> DNA  
<213> Homo sapiens

<400> 3  
atgctggcag ctgcctttgc agactctaac tccagcagca tgaatgtgtc ctttgctcac 60  
ctccactttg ccggagggtta cctgccctct gattcccagg actggagaac catcatcccg 120  
gctctcttgg tggtgtgtct cctggtgggc ttcgtgggaa acctgtgtgt gattggcatc 180  
ctccttcaca atgcttggaaggaaagcca tccatgatcc actccctgat tctgaatctc 240  
agcctggctg atctctccct cctgctgttt tctgcacctt tccgagctac ggcgtactcc 300  
aaaagtgttt gggatctagg ctggtttgtc tgcaagtcct ctgactgggt tatccacaca 360  
tgcattggcag ccaagagcct gacaatcgtt gtggtggcca aagtatgctt catgtatgca 420  
agtgaccag ccaagcaagt gaggatccac aactacacca tctggtcagt gctggtggcc 480  
atctggactg tggctagcct gttacccctg ccggaatggt tctttagcac catcaggcat 540  
catgaagggtg tggaaatgtg cctcgtggat gtaccagctg tggctgaaga gtttatgtcg 600  
atgttttgta agctctaccc actcctggca tttggccttc cattattttt tgccagcttt 660  
tatttctgga gagcttatga ccaatgtaaa aaacgaggaa ctaagactca aaatcttaga 720  
aaccagatac gctcaaagca agtcacagtg atgtgctga gcattgccat catctctgct 780  
ctcttgtggc tccccgaatg ggtagcttgg ctgtgggtat ggcatctgaa ggctgcaggc 840  
ccggccccac cacaagggtt catagccctg tctcaagtct tgatgttttc catctcttca 900  
gcaaaccctc tcatttttct tgtgatgtcg gaagagttca gggaaggctt gaaagggtga 960  
tggaatgga tgataaccaa aaaacctcca actgtctcag agtctcagga aacaccagct 1020  
ggcaactcag aggtctctcc tgacaagggt ccatctccag aatccccagc atccatacca 1080  
gaaaaagaga aaccagctc tccctcctct ggcaaggga aaactgagaa ggcagagatt 1140  
cccatccttc ctgacgtaga gcagttttgg catgagaggg acacagtccc ttctgtacag 1200  
gacaatgacc ctatcccctg ggaacatgaa gatcaagaga cagggaagg tgttaaatag 1260

<210> 4  
<211> 419  
<212> PRT  
<213> Homo sapiens

&lt;400&gt; 4

```

Met Leu Ala Ala Ala Phe Ala Asp Ser Asn Ser Ser Ser Met Asn Val
1          5          10          15

Ser Phe Ala His Leu His Phe Ala Gly Gly Tyr Leu Pro Ser Asp Ser
20          25          30

Gln Asp Trp Arg Thr Ile Ile Pro Ala Leu Leu Val Ala Val Cys Leu
35          40          45

Val Gly Phe Val Gly Asn Leu Cys Val Ile Gly Ile Leu Leu His Asn
50          55          60

Ala Trp Lys Gly Lys Pro Ser Met Ile His Ser Leu Ile Leu Asn Leu
65          70          75          80

Ser Leu Ala Asp Leu Ser Leu Leu Leu Phe Ser Ala Pro Ile Arg Ala
85          90          95

Thr Ala Tyr Ser Lys Ser Val Trp Asp Leu Gly Trp Phe Val Cys Lys
100         105         110

Ser Ser Asp Trp Phe Ile His Thr Cys Met Ala Ala Lys Ser Leu Thr
115         120         125

Ile Val Val Val Ala Lys Val Cys Phe Met Tyr Ala Ser Asp Pro Ala
130         135         140

Lys Gln Val Ser Ile His Asn Tyr Thr Ile Trp Ser Val Leu Val Ala
145         150         155         160

Ile Trp Thr Val Ala Ser Leu Leu Pro Leu Pro Glu Trp Phe Phe Ser
165         170         175

Thr Ile Arg His His Glu Gly Val Glu Met Cys Leu Val Asp Val Pro
180         185         190

Ala Val Ala Glu Glu Phe Met Ser Met Phe Gly Lys Leu Tyr Pro Leu
195         200         205

Leu Ala Phe Gly Leu Pro Leu Phe Phe Ala Ser Phe Tyr Phe Trp Arg
210         215         220

Ala Tyr Asp Gln Cys Lys Lys Arg Gly Thr Lys Thr Gln Asn Leu Arg
225         230         235         240

Asn Gln Ile Arg Ser Lys Gln Val Thr Val Met Leu Leu Ser Ile Ala
245         250         255

Ile Ile Ser Ala Leu Leu Trp Leu Pro Glu Trp Val Ala Trp Leu Trp
260         265         270

Val Trp His Leu Lys Ala Ala Gly Pro Ala Pro Pro Gln Gly Phe Ile
275         280         285

Ala Leu Ser Gln Val Leu Met Phe Ser Ile Ser Ser Ala Asn Pro Leu
290         295         300

Ile Phe Leu Val Met Ser Glu Glu Phe Arg Glu Gly Leu Lys Gly Val
305         310         315         320

Trp Lys Trp Met Ile Thr Lys Lys Pro Pro Thr Val Ser Glu Ser Gln
325         330         335

```

Glu Thr Pro Ala Gly Asn Ser Glu Gly Leu Pro Asp Lys Val Pro Ser  
                   340                  345                  350  
 Pro Glu Ser Pro Ala Ser Ile Pro Glu Lys Glu Lys Pro Ser Ser Pro  
                   355                  360                  365  
 Ser Ser Gly Lys Gly Lys Thr Glu Lys Ala Glu Ile Pro Ile Leu Pro  
                   370                  375                  380  
 Asp Val Glu Gln Phe Trp His Glu Arg Asp Thr Val Pro Ser Val Gln  
                   385                  390                  395                  400  
 Asp Asn Asp Pro Ile Pro Trp Glu His Glu Asp Gln Glu Thr Gly Glu  
                   405                  410                  415

Gly Val Lys

<210> 5  
 <211> 1014  
 <212> DNA  
 <213> Homo sapiens

<400> 5  
 atggggaacg attctgtcag ctacgagtat ggggattaca ggcacctctc ggaccgccct 60  
 gtggactgcc tggatggcgc ctgcctggcc atcgaccgcg tgcgcgtggc cccgctccca 120  
 ctgtatgccg ccatcttcct ggtgggggtg ccgggcaatg ccatggtggc ctgggtggct 180  
 gggaagggtg cccgccggag ggtgggtgcc acctggttgc tccacctggc cgtggcggat 240  
 ttgctgtgct gtttgtctct gcccatcctg gcagtgccca ttgccctgg aggccactgg 300  
 ccgtatggtg cagtgggctg tcgggcgctg ccctccatca tcctgctgac catgtatgcc 360  
 agcgtcctgc tcctggcagc tctcagtgcc gacctctgct tcctggctct cgggcctgcc 420  
 tgggtggtcta cggttcagcg ggcgtgcggg gtgcaggtgg cctgtggggc agcctggaca 480  
 ctggccttgc tgetcaccgt gccctccgcc atctaccgcc ggctgcacca ggagcacttc 540  
 ccagcccggc tgcagtgtgt ggtggactac ggcggctcct ccagcaccga gaatgcggtg 600  
 actgccatcc ggtttctttt tggcttcctg gggcccttgg tggccgtggc cagctgccac 660  
 agtgccctcc tgtgtgggc agcccagcgc tgccggccgc tgggcacagc cattgtggtg 720  
 gggttttttg tctgtgggc accctaccac ctgctggggc tgggtgtcac tgtggcggcc 780  
 ccgaactccg cactcctggc cagggccctg cgggctgaac ccctcatcgt gggccttgcc 840  
 ctcgtcaca getgcctcaa tccatgctc ttctgtatt ttgggagggc tcaactccgc 900  
 cggtcactgc cagctgcctg tcactgggcc ctgagggagt cccagggcca ggacgaaagt 960  
 gtggacagca agaaatccac cagccatgac ctggtctcgg agatggaggt gtag 1014

<210> 6  
 <211> 337  
 <212> PRT  
 <213> Homo sapiens

&lt;400&gt; 6

```

Met Gly Asn Asp Ser Val Ser Tyr Glu Tyr Gly Asp Tyr Ser Asp Leu
1          5          10          15

Ser Asp Arg Pro Val Asp Cys Leu Asp Gly Ala Cys Leu Ala Ile Asp
20          25          30

Pro Leu Arg Val Ala Pro Leu Pro Leu Tyr Ala Ala Ile Phe Leu Val
35          40          45

Gly Val Pro Gly Asn Ala Met Val Ala Trp Val Ala Gly Lys Val Ala
50          55          60

Arg Arg Arg Val Gly Ala Thr Trp Leu Leu His Leu Ala Val Ala Asp
65          70          75          80

Leu Leu Cys Cys Leu Ser Leu Pro Ile Leu Ala Val Pro Ile Ala Arg
85          90          95

Gly Gly His Trp Pro Tyr Gly Ala Val Gly Cys Arg Ala Leu Pro Ser
100         105         110

Ile Ile Leu Leu Thr Met Tyr Ala Ser Val Leu Leu Leu Ala Ala Leu
115         120         125

Ser Ala Asp Leu Cys Phe Leu Ala Leu Gly Pro Ala Trp Trp Ser Thr
130         135         140

Val Gln Arg Ala Cys Gly Val Gln Val Ala Cys Gly Ala Ala Trp Thr
145         150         155         160

Leu Ala Leu Leu Leu Thr Val Pro Ser Ala Ile Tyr Arg Arg Leu His
165         170         175

Gln Glu His Phe Pro Ala Arg Leu Gln Cys Val Val Asp Tyr Gly Gly
180         185         190

Ser Ser Ser Thr Glu Asn Ala Val Thr Ala Ile Arg Phe Leu Phe Gly
195         200         205

Phe Leu Gly Pro Leu Val Ala Val Ala Ser Cys His Ser Ala Leu Leu
210         215         220

Cys Trp Ala Ala Arg Arg Cys Arg Pro Leu Gly Thr Ala Ile Val Val
225         230         235         240

Gly Phe Phe Val Cys Trp Ala Pro Tyr His Leu Leu Gly Leu Val Leu
245         250         255

Thr Val Ala Ala Pro Asn Ser Ala Leu Leu Ala Arg Ala Leu Arg Ala
260         265         270

Glu Pro Leu Ile Val Gly Leu Ala Leu Ala His Ser Cys Leu Asn Pro
275         280         285

Met Leu Phe Leu Tyr Phe Gly Arg Ala Gln Leu Arg Arg Ser Leu Pro
290         295         300

Ala Ala Cys His Trp Ala Leu Arg Glu Ser Gln Gly Gln Asp Glu Ser
305         310         315         320

Val Asp Ser Lys Lys Ser Thr Ser His Asp Leu Val Ser Glu Met Glu
325         330         335

```

Val

<210> 7  
 <211> 1272  
 <212> DNA  
 <213> Homo sapiens

<400> 7  
 atgttggtgc accgtggtgg ccagctgata gtgccaatca tcccactttg ccctgagcac 60  
 tcctgcaggg gtagaagact ccagaacctt ctctcaggcc catggcccaa gcagcccatg 120  
 gaacttcata acctgagctc tccatctccc tctctctcct cctctgttct ccctccctcc 180  
 ttctctccct caccctctc tgctccctct gcctttacca ctgtggggg gtcctctgga 240  
 gggccctgcc accccacctc ttcctcgctg gtgtctgcct tcctggcacc aatcctggcc 300  
 ctggagtttg tcctgggcct ggtggggaac agtttgggcc tcttcattct ctgcatccac 360  
 acgcggccct ggacctcaa cacggtgttc ctggctcagcc tgggtggcgc tgacttcctc 420  
 ctgatcagca acctgcccct ccgcgtggac tactacctcc tccatgagac ctggcgcttt 480  
 ggggctgctg cctgcaaagt caacctcttc atgctgtcca ccaaccgcac ggccagcggt 540  
 gtcttctcct cagccatcgc actcaaccgc tacctgaagg tgggtgcagcc ccaccacgtg 600  
 ctgagccgtg ctcccggtgg ggcagctgcc cgggtggcgc ggggactctg ggtgggcatc 660  
 ctgctcctca acgggcacct gctcctgagc acctctctcg gccctcctg cctcagctac 720  
 aggggtggga cgaagccctc ggctcgtc cgctggcacc aggcactgta cctgctggag 780  
 ttcttctcgc cactggcgct catcctcttt gctattgtga gcattgggct caccatccgg 840  
 aaccgtggtc tgggcgggca ggcaggcccg cagagggcca tgcgtgtgct ggccatggtg 900  
 gtggccgtct acaccatctg cttcttgccc agcatcatct ttggcatggc ttccatggtg 960  
 gctttctggc tgtccgcctg ccgatccctg gacctctgca cacagctctt ccatggctcc 1020  
 ctggccttca cctacctcaa cagtgtcctg gaccccgctc tctactgctt ctctagcccc 1080  
 aacttcctcc accagagccg ggccttgctg ggcctcacgc ggggccgga gggccagtg 1140  
 agcgacgaga gtcctacca accctccagg cagtggcgct accgggaggc ctctaggaag 1200  
 gcggaggcca tagggaagct gaaagtgcag ggcgaggtct ctctggaaaa ggaaggctcc 1260  
 tcccagggt ga 1272

<210> 8  
 <211> 423  
 <212> PRT  
 <213> Homo sapiens

<400> 8

Met Leu Cys His Arg Gly Gly Gln Leu Ile Val Pro Ile Ile Pro Leu  
 1 5 10 15

Cys Pro Glu His Ser Cys Arg Gly Arg Arg Leu Gln Asn Leu Leu Ser

20										25										30										
Gly	Pro	Trp	Pro	Lys	Gln	Pro	Met	Glu	Leu	His	Asn	Leu	Ser	Ser	Pro															
		35					40					45																		
Ser	Pro	Ser	Leu	Ser	Ser	Ser	Val	Leu	Pro	Pro	Ser	Phe	Ser	Pro	Ser															
		50				55					60																			
Pro	Ser	Ser	Ala	Pro	Ser	Ala	Phe	Thr	Thr	Val	Gly	Gly	Ser	Ser	Gly															
		65			70					75					80															
Gly	Pro	Cys	His	Pro	Thr	Ser	Ser	Ser	Leu	Val	Ser	Ala	Phe	Leu	Ala															
				85					90					95																
Pro	Ile	Leu	Ala	Leu	Glu	Phe	Val	Leu	Gly	Leu	Val	Gly	Asn	Ser	Leu															
			100					105					110																	
Ala	Leu	Phe	Ile	Phe	Cys	Ile	His	Thr	Arg	Pro	Trp	Thr	Ser	Asn	Thr															
		115				120						125																		
Val	Phe	Leu	Val	Ser	Leu	Val	Ala	Ala	Asp	Phe	Leu	Leu	Ile	Ser	Asn															
		130				135					140																			
Leu	Pro	Leu	Arg	Val	Asp	Tyr	Tyr	Leu	Leu	His	Glu	Thr	Trp	Arg	Phe															
		145			150					155					160															
Gly	Ala	Ala	Ala	Cys	Lys	Val	Asn	Leu	Phe	Met	Leu	Ser	Thr	Asn	Arg															
				165					170					175																
Thr	Ala	Ser	Val	Val	Phe	Leu	Thr	Ala	Ile	Ala	Leu	Asn	Arg	Tyr	Leu															
			180					185					190																	
Lys	Val	Val	Gln	Pro	His	His	Val	Leu	Ser	Arg	Ala	Ser	Val	Gly	Ala															
		195					200						205																	
Ala	Ala	Arg	Val	Ala	Gly	Gly	Leu	Trp	Val	Gly	Ile	Leu	Leu	Leu	Asn															
		210				215					220																			
Gly	His	Leu	Leu	Leu	Ser	Thr	Phe	Ser	Gly	Pro	Ser	Cys	Leu	Ser	Tyr															
		225			230				235					240																
Arg	Val	Gly	Thr	Lys	Pro	Ser	Ala	Ser	Leu	Arg	Trp	His	Gln	Ala	Leu															
				245					250					255																
Tyr	Leu	Leu	Glu	Phe	Phe	Leu	Pro	Leu	Ala	Leu	Ile	Leu	Phe	Ala	Ile															
			260					265					270																	
Val	Ser	Ile	Gly	Leu	Thr	Ile	Arg	Asn	Arg	Gly	Leu	Gly	Gly	Gln	Ala															
		275					280					285																		
Gly	Pro	Gln	Arg	Ala	Met	Arg	Val	Leu	Ala	Met	Val	Val	Ala	Val	Tyr															
		290				295					300																			
Thr	Ile	Cys	Phe	Leu	Pro	Ser	Ile	Ile	Phe	Gly	Met	Ala	Ser	Met	Val															
		305			310					315				320																
Ala	Phe	Trp	Leu	Ser	Ala	Cys	Arg	Ser	Leu	Asp	Leu	Cys	Thr	Gln	Leu															
				325					330					335																
Phe	His	Gly	Ser	Leu	Ala	Phe	Thr	Tyr	Leu	Asn	Ser	Val	Leu	Asp	Pro															
		340						345					350																	
Val	Leu	Tyr	Cys	Phe	Ser	Ser	Pro	Asn	Phe	Leu	His	Gln	Ser	Arg	Ala															
		355					360					365																		

Leu Leu Gly Leu Thr Arg Gly Arg Gln Gly Pro Val Ser Asp Glu Ser  
 370 375 380

Ser Tyr Gln Pro Ser Arg Gln Trp Arg Tyr Arg Glu Ala Ser Arg Lys  
 385 390 395 400

Ala Glu Ala Ile Gly Lys Leu Lys Val Gln Gly Glu Val Ser Leu Glu  
 405 410 415

Lys Glu Gly Ser Ser Gln Gly  
 420

<210> 9  
 <211> 966  
 <212> DNA  
 <213> Homo sapiens

<400> 9  
 atgaaccaga ctttgaatag cagtgggacc gtggagtcag ccctaaacta ttccagaggg 60  
 agcacagtgc acacggccta cctggtgctg agctccctgg ccatgttcac ctgcctgtgc 120  
 gggatggcag gcaacagcat ggtgatctgg ctgctgggct ttcgaatgca caggaacccc 180  
 ttctgcatct atatcctcaa cctggcggca gccgacctcc tcttcctctt cagcatggct 240  
 tccacgctca gcctggaac ccagccctg gtcaatacca ctgacaaggc ccacgagctg 300  
 atgaagagac tgatgtactt tgcctacaca gtgggcctga gcctgctgac ggccatcagc 360  
 acccagcgct gtctctctgt cctcttcctt atctggttca agtgtcacgc gcccaggcac 420  
 ctgtcagcct ggggtgtgtg cctgctgtgg aactctgtc tcctgatgaa cgggttgacc 480  
 tcttccttct gcagcaagtt cttgaaattc aatgaagatc ggtgcttcag ggtggacatg 540  
 gtccaggccg ccctcatcat gggggtctta acccagtgta tgactctgtc cagcctgacc 600  
 ctctttgtct ggggtgcggag gagctcccag cagtggcggc ggagccccc acggctgttc 660  
 gtggtggtcc tggcctctgt cctggtgttc ctcatctgtt ccctgcctct gagcatctac 720  
 tgggttgtgc tctactggtt gagcctgccc cccgagatgc aggtcctgtg cttcagcttg 780  
 tcacgcctct cctcgtccgt aagcagcagc gccaaacccc tcatctactt cctggtgggc 840  
 agccggagga gccacaggct gccaccagg tccctgggga ctgtgctcca acaggcgctt 900  
 cgcgaggagc ccgagctgga aggtggggag acgcccaccg tgggcaccaa tgagatgggg 960  
 gcttga 966

<210> 10  
 <211> 321  
 <212> PRT  
 <213> Homo sapiens

<400> 10

Met Asn Gln Thr Leu Asn Ser Ser Gly Thr Val Glu Ser Ala Leu Asn  
 1 5 10 15

Tyr Ser Arg Gly Ser Thr Val His Thr Ala Tyr Leu Val Leu Ser Ser  
 20 25 30

Leu Ala Met Phe Thr Cys Leu Cys Gly Met Ala Gly Asn Ser Met Val  
 35 40 45  
 Ile Trp Leu Leu Gly Phe Arg Met His Arg Asn Pro Phe Cys Ile Tyr  
 50 55 60  
 Ile Leu Asn Leu Ala Ala Ala Asp Leu Leu Phe Leu Phe Ser Met Ala  
 65 70 75 80  
 Ser Thr Leu Ser Leu Glu Thr Gln Pro Leu Val Asn Thr Thr Asp Lys  
 85 90 95  
 Val His Glu Leu Met Lys Arg Leu Met Tyr Phe Ala Tyr Thr Val Gly  
 100 105 110  
 Leu Ser Leu Leu Thr Ala Ile Ser Thr Gln Arg Cys Leu Ser Val Leu  
 115 120 125  
 Phe Pro Ile Trp Phe Lys Cys His Arg Pro Arg His Leu Ser Ala Trp  
 130 135 140  
 Val Cys Gly Leu Leu Trp Thr Leu Cys Leu Leu Met Asn Gly Leu Thr  
 145 150 155 160  
 Ser Ser Phe Cys Ser Lys Phe Leu Lys Phe Asn Glu Asp Arg Cys Phe  
 165 170 175  
 Arg Val Asp Met Val Gln Ala Ala Leu Ile Met Gly Val Leu Thr Pro  
 180 185 190  
 Val Met Thr Leu Ser Ser Leu Thr Leu Phe Val Trp Val Arg Arg Ser  
 195 200 205  
 Ser Gln Gln Trp Arg Arg Gln Pro Thr Arg Leu Phe Val Val Val Leu  
 210 215 220  
 Ala Ser Val Leu Val Phe Leu Ile Cys Ser Leu Pro Leu Ser Ile Tyr  
 225 230 235 240  
 Trp Phe Val Leu Tyr Trp Leu Ser Leu Pro Pro Glu Met Gln Val Leu  
 245 250 255  
 Cys Phe Ser Leu Ser Arg Leu Ser Ser Ser Val Ser Ser Ser Ala Asn  
 260 265 270  
 Pro Val Ile Tyr Phe Leu Val Gly Ser Arg Arg Ser His Arg Leu Pro  
 275 280 285  
 Thr Arg Ser Leu Gly Thr Val Leu Gln Gln Ala Leu Arg Glu Glu Pro  
 290 295 300  
 Glu Leu Glu Gly Gly Glu Thr Pro Thr Val Gly Thr Asn Glu Met Gly  
 305 310 315 320

Ala

<210> 11  
 <211> 1356  
 <212> DNA  
 <213> Homo sapiens

<400> 11  
 atggagtcct caccatccc ccagtcacatca gggaactctt ccactttggg gagggtcctt

60



```

caaacccag gtccctctac tgccagtggg gtcccgagg tggggctacg ggatgttgct 120
tcggaatctg tggccctctt cttcatgctc ctgctggact tgactgctgt ggctggcaat 180
gccgctgtga tggccgtgat cgccaagacg cctgccctcc gaaaatttgt cttcgtcttc 240
cacctctgcc tgggtggacct gctggctgcc ctgaccctca tgcccctggc catgctctcc 300
agctctgccc tctttgacca cgccctcttt ggggaggtgg cctgccgcct ctacttgttt 360
ctgagcgtgt gctttgtcag cctggccatc ctctcgtgt cagccatcaa tgtggagcgc 420
tactattacg tagtccaccc catgcgctac gaggtgcga tgacgctggg gctggtggcc 480
tctgtgctgg tgggtgtgtg ggtgaaggcc ttggccatgg cttctgtgcc agtgttggga 540
agggtctcct gggaggaagg agctcccagt gtccccccag gctgttcaact ccagtggagc 600
cacagtgcct actgccagct ttttgtggtg gtctttgctg tcctttactt tctgttgccc 660
ctgctcctca tacttggtgt ctactgcagc atgttccgag tggcccgct ggctgccatg 720
cagcacgggc cgctgccac gtggatggag acacccggc aacgctccga atctctcagc 780
agccgtcca cgatggtcac cagctcgggg gccccccaga ccacccaca ccggacgttt 840
gggggaggga aagcagcagt ggttctcctg gctgtggggg gacagttcct gctctgttgg 900
ttgccctact tctctttcca cctctatgtt gccctgagt ctcagcccat ttcaactggg 960
cagggtgaga gtgtggtcac ctggattggc tacttttgct tcacttcaa ccctttcttc 1020
tatggatgtc tcaaccggca gatccgggg gagctcagca agcagtttgt ctgcttcttc 1080
aagccagctc cagaggagga gctgaggctg cctagccggg agggctccat tgaggagaac 1140
ttcctgcagt tccttcaggg gactggctgt cttctgagt cctgggtttc ccgacccta 1200
cccagcccca agcaggagcc acctgctgtt gactttcgaa tcccaggcca gatagctgag 1260
gagacctctg agttcctgga gcagcaactc accagcgaca tcatcatgtc agacagctac 1320
ctccgtcctg ccgcctcacc ccggctggag tcatga 1356

```

```

<210> 12
<211> 451
<212> PRT
<213> Homo sapiens

```

```

<400> 12

```

```

Met Glu Ser Ser Pro Ile Pro Gln Ser Ser Gly Asn Ser Ser Thr Leu
1           5           10           15

```

```

Gly Arg Val Pro Gln Thr Pro Gly Pro Ser Thr Ala Ser Gly Val Pro
20           25           30

```

```

Glu Val Gly Leu Arg Asp Val Ala Ser Glu Ser Val Ala Leu Phe Phe
35           40           45

```

```

Met Leu Leu Leu Asp Leu Thr Ala Val Ala Gly Asn Ala Ala Val Met
50           55           60

```

Ala Val Ile Ala Lys Thr Pro Ala Leu Arg Lys Phe Val Phe Val Phe  
 65 70 75 80  
 His Leu Cys Leu Val Asp Leu Leu Ala Ala Leu Thr Leu Met Pro Leu  
 85 90 95  
 Ala Met Leu Ser Ser Ser Ala Leu Phe Asp His Ala Leu Phe Gly Glu  
 100 105 110  
 Val Ala Cys Arg Leu Tyr Leu Phe Leu Ser Val Cys Phe Val Ser Leu  
 115 120 125  
 Ala Ile Leu Ser Val Ser Ala Ile Asn Val Glu Arg Tyr Tyr Tyr Val  
 130 135 140  
 Val His Pro Met Arg Tyr Glu Val Arg Met Thr Leu Gly Leu Val Ala  
 145 150 155 160  
 Ser Val Leu Val Gly Val Trp Val Lys Ala Leu Ala Met Ala Ser Val  
 165 170 175  
 Pro Val Leu Gly Arg Val Ser Trp Glu Glu Gly Ala Pro Ser Val Pro  
 180 185 190  
 Pro Gly Cys Ser Leu Gln Trp Ser His Ser Ala Tyr Cys Gln Leu Phe  
 195 200 205  
 Val Val Val Phe Ala Val Leu Tyr Phe Leu Leu Pro Leu Leu Leu Ile  
 210 215 220  
 Leu Val Val Tyr Cys Ser Met Phe Arg Val Ala Arg Val Ala Ala Met  
 225 230 235 240  
 Gln His Gly Pro Leu Pro Thr Trp Met Glu Thr Pro Arg Gln Arg Ser  
 245 250 255  
 Glu Ser Leu Ser Ser Arg Ser Thr Met Val Thr Ser Ser Gly Ala Pro  
 260 265 270  
 Gln Thr Thr Pro His Arg Thr Phe Gly Gly Gly Lys Ala Ala Val Val  
 275 280 285  
 Leu Leu Ala Val Gly Gly Gln Phe Leu Leu Cys Trp Leu Pro Tyr Phe  
 290 295 300  
 Ser Phe His Leu Tyr Val Ala Leu Ser Ala Gln Pro Ile Ser Thr Gly  
 305 310 315 320  
 Gln Val Glu Ser Val Val Thr Trp Ile Gly Tyr Phe Cys Phe Thr Ser  
 325 330 335  
 Asn Pro Phe Phe Tyr Gly Cys Leu Asn Arg Gln Ile Arg Gly Glu Leu  
 340 345 350  
 Ser Lys Gln Phe Val Cys Phe Phe Lys Pro Ala Pro Glu Glu Glu Leu  
 355 360 365  
 Arg Leu Pro Ser Arg Glu Gly Ser Ile Glu Glu Asn Phe Leu Gln Phe  
 370 375 380  
 Leu Gln Gly Thr Gly Cys Pro Ser Glu Ser Trp Val Ser Arg Pro Leu  
 385 390 395 400  
 Pro Ser Pro Lys Gln Glu Pro Pro Ala Val Asp Phe Arg Ile Pro Gly  
 405 410 415

Gln Ile Ala Glu Glu Thr Ser Glu Phe Leu Glu Gln Gln Leu Thr Ser  
420 425 430

Asp Ile Ile Met Ser Asp Ser Tyr Leu Arg Pro Ala Ala Ser Pro Arg  
435 440 445

Leu Glu Ser  
450

<210> 13  
<211> 1041  
<212> DNA  
<213> Homo sapiens

<400> 13  
atggagagaa aatttatgtc cttgcaacca tccatctccg tatcagaaat ggaaccaa 60  
ggcaccttca gcaataacaa cagcaggaac tgcacaattg aaaacttcaa gagagaattt 120  
ttccaattg tatatctgat aatatttttc tggggagtct tgggaaatgg gttgtccata 180  
tatgttttcc tgcagcctta taagaagtcc acatctgtga acgttttcat gctaaatctg 240  
gccatttcag atctcctggt cataagcacg cttcccttca gggctgacta ttatcttaga 300  
ggctccaatt ggatatttgg agacctggcc tgcaggatta tgtcttattc cttgtatgtc 360  
aacatgtaca gcagtattta tttcctgacc gtgctgagtg ttgtgcgttt cctggcaatg 420  
gttcacccct ttcggcttct gcatgtcacc agcatcagga gtgcctggat cctctgtggg 480  
atcatatgga tccttatcat ggcttcctca ataatgctcc tggacagtgg ctctgagcag 540  
aacggcagtg tcacatcatg cttagagctg aatctctata aaattgctaa gctgcagacc 600  
atgaactata ttgccttggg ggtgggctgc ctgctgccat tttcacact cagcatctgt 660  
tatctgctga tcattcgggt tctgttaaaa gtggaggtcc cagaatcggg gctgcgggtt 720  
tctcacagga aggcactgac caccatcatc atcaccttga tcatcttctt cttgtgtttc 780  
ctgccctatc acacactgag gaccgtccac ttgacgacat ggaaagtggg tttatgcaaa 840  
gacagactgc ataaagcttt ggttatcaca ctggccttgg cagcagccaa tgcctgcttc 900  
aatcctctgc tctattactt tgctggggag aattttaagg acagactaaa gtctgcactc 960  
agaaaaggcc atccacagaa ggcaaagaca aagtgtgttt tccctgttag tgtgtggttg 1020  
agaaaggaaa caagagtata a 1041

<210> 14  
<211> 346  
<212> PRT  
<213> Homo sapiens

<400> 14

Met Glu Arg Lys Phe Met Ser Leu Gln Pro Ser Ile Ser Val Ser Glu  
1 5 10 15

Met Glu Pro Asn Gly Thr Phe Ser Asn Asn Ser Arg Asn Cys Thr  
20 25 30

Ile Glu Asn Phe Lys Arg Glu Phe Phe Pro Ile Val Tyr Leu Ile Ile  
 35 40 45  
 Phe Phe Trp Gly Val Leu Gly Asn Gly Leu Ser Ile Tyr Val Phe Leu  
 50 55 60  
 Gln Pro Tyr Lys Lys Ser Thr Ser Val Asn Val Phe Met Leu Asn Leu  
 65 70 75 80  
 Ala Ile Ser Asp Leu Leu Phe Ile Ser Thr Leu Pro Phe Arg Ala Asp  
 85 90 95  
 Tyr Tyr Leu Arg Gly Ser Asn Trp Ile Phe Gly Asp Leu Ala Cys Arg  
 100 105 110  
 Ile Met Ser Tyr Ser Leu Tyr Val Asn Met Tyr Ser Ser Ile Tyr Phe  
 115 120 125  
 Leu Thr Val Leu Ser Val Val Arg Phe Leu Ala Met Val His Pro Phe  
 130 135 140  
 Arg Leu Leu His Val Thr Ser Ile Arg Ser Ala Trp Ile Leu Cys Gly  
 145 150 155 160  
 Ile Ile Trp Ile Leu Ile Met Ala Ser Ser Ile Met Leu Leu Asp Ser  
 165 170 175  
 Gly Ser Glu Gln Asn Gly Ser Val Thr Ser Cys Leu Glu Leu Asn Leu  
 180 185 190  
 Tyr Lys Ile Ala Lys Leu Gln Thr Met Asn Tyr Ile Ala Leu Val Val  
 195 200 205  
 Gly Cys Leu Leu Pro Phe Phe Thr Leu Ser Ile Cys Tyr Leu Leu Ile  
 210 215 220  
 Ile Arg Val Leu Leu Lys Val Glu Val Pro Glu Ser Gly Leu Arg Val  
 225 230 235 240  
 Ser His Arg Lys Ala Leu Thr Thr Ile Ile Ile Thr Leu Ile Ile Phe  
 245 250 255  
 Phe Leu Cys Phe Leu Pro Tyr His Thr Leu Arg Thr Val His Leu Thr  
 260 265 270  
 Thr Trp Lys Val Gly Leu Cys Lys Asp Arg Leu His Lys Ala Leu Val  
 275 280 285  
 Ile Thr Leu Ala Leu Ala Ala Asn Ala Cys Phe Asn Pro Leu Leu  
 290 295 300  
 Tyr Tyr Phe Ala Gly Glu Asn Phe Lys Asp Arg Leu Lys Ser Ala Leu  
 305 310 315 320  
 Arg Lys Gly His Pro Gln Lys Ala Lys Thr Lys Cys Val Phe Pro Val  
 325 330 335  
 Ser Val Trp Leu Arg Lys Glu Thr Arg Val  
 340 345

<210> 15  
 <211> 1527  
 <212> DNA  
 <213> Homo sapiens

<400> 15  
 atgacgtcca cctgcaccaa cagcacgcgc gagagtaaca gcagccacac gtgcatgccc 60  
 ctctccaaaa tgcccatcag cctggcccac ggcacatcc gctcaaccgt gctgggtatc 120  
 ttcctcgccg cctctttcgt cggcaacata gtgctggcgc tagtggtgca gcgcaagccg 180  
 cagctgctgc aggtgaccaa ccgttttata tttaacctcc tcgtcaccga cctgctgcag 240  
 atttcgctcg tggccccctg ggtggtggcc acctctgtgc ctctcttctg gcccccaac 300  
 agccacttct gcacggccct ggtagcctc acccacctgt tcgccttcgc cagcgtcaac 360  
 accattgtcg tgggtgtcagt ggatcgctac ttgtccatca tccaccctct ctctacccg 420  
 tccaagatga cccagcgccg cggttacctg ctctctatg gcacctggat tgtggccatc 480  
 ctgcagagca ctctccact ctacggctgg ggcaggctg cctttgatga gcgcaatgct 540  
 ctctgctcca tgatctgggg ggccagcccc agctacacta ttctcagcgt ggtgtccttc 600  
 atcgtcattc cactgattgt catgattgcc tgctactccg tgggtgtctg tgcagcccgg 660  
 aggcagcatg ctctgctgta caatgtcaag agacacagct tggaagtgcg agtcaaggac 720  
 tgttgggaga atgaggatga agaggagca gagaagaagg aggagtcca ggatgagagt 780  
 gagtttcgcc gccagcatga aggtgaggtc aaggccaagg agggcagaat ggaagccaag 840  
 gacggcagcc tgaaggccaa ggaaggaagc acggggacca gtgagagtag ttagaggcc 900  
 aggggcagcg aggaggtcag agagagcagc acggtggcca gcgacggcag catggagggt 960  
 aaggaaggca gcaccaaagt tgaggagaac agcatgaagg cagacaaggg tcgcacagag 1020  
 gtcaaccagt gcagcattga cttgggtgaa gatgacatgg agtttgggtga agacgacatc 1080  
 aatttcagtg aggatgacgt cgaggcagtg aacatcccgg agagcctccc acccagtcgt 1140  
 cgtaacagca acagcaaccc tcctctgccc aggtgctacc agtgcaaagc tgctaaagtg 1200  
 atcttcatca tcattttctc ctatgtgcta tccctggggc cctactgctt tttagcagtc 1260  
 ctggccgtgt ggggtgatgt cgaaacccag gtaccccagt gggatgacac cataatcatc 1320  
 tggcttttct tcctgcagtg ctgcatccac ccctatgtct atggctacat gcacaagacc 1380  
 attaagaagg aaatccagga catgctgaag aagttcttct gcaaggaaaa gccccgaaa 1440  
 gaagatagcc acccagacct gcccgaaca gaggggtggga ctgaaggcaa gattgtccct 1500  
 tcctacgatt ctgtacttt tccttga 1527

<210> 16  
 <211> 508  
 <212> PRT  
 <213> Homo sapiens

<400> 16

Met Thr Ser Thr Cys Thr Asn Ser Thr Arg Glu Ser Asn Ser Ser His  
 1 5 10 15

Thr Cys Met Pro Leu Ser Lys Met Pro Ile Ser Leu Ala His Gly Ile  
 20 25 30  
 Ile Arg Ser Thr Val Leu Val Ile Phe Leu Ala Ala Ser Phe Val Gly  
 35 40 45  
 Asn Ile Val Leu Ala Leu Val Leu Gln Arg Lys Pro Gln Leu Leu Gln  
 50 55 60  
 Val Thr Asn Arg Phe Ile Phe Asn Leu Leu Val Thr Asp Leu Leu Gln  
 65 70 75 80  
 Ile Ser Leu Val Ala Pro Trp Val Val Ala Thr Ser Val Pro Leu Phe  
 85 90 95  
 Trp Pro Leu Asn Ser His Phe Cys Thr Ala Leu Val Ser Leu Thr His  
 100 105 110  
 Leu Phe Ala Phe Ala Ser Val Asn Thr Ile Val Val Val Ser Val Asp  
 115 120 125  
 Arg Tyr Leu Ser Ile Ile His Pro Leu Ser Tyr Pro Ser Lys Met Thr  
 130 135 140  
 Gln Arg Arg Gly Tyr Leu Leu Leu Tyr Gly Thr Trp Ile Val Ala Ile  
 145 150 155 160  
 Leu Gln Ser Thr Pro Pro Leu Tyr Gly Trp Gly Gln Ala Ala Phe Asp  
 165 170 175  
 Glu Arg Asn Ala Leu Cys Ser Met Ile Trp Gly Ala Ser Pro Ser Tyr  
 180 185 190  
 Thr Ile Leu Ser Val Val Ser Phe Ile Val Ile Pro Leu Ile Val Met  
 195 200 205  
 Ile Ala Cys Tyr Ser Val Val Phe Cys Ala Ala Arg Arg Gln His Ala  
 210 215 220  
 Leu Leu Tyr Asn Val Lys Arg His Ser Leu Glu Val Arg Val Lys Asp  
 225 230 235 240  
 Cys Val Glu Asn Glu Asp Glu Glu Gly Ala Glu Lys Lys Glu Glu Phe  
 245 250 255  
 Gln Asp Glu Ser Glu Phe Arg Arg Gln His Glu Gly Glu Val Lys Ala  
 260 265 270  
 Lys Glu Gly Arg Met Glu Ala Lys Asp Gly Ser Leu Lys Ala Lys Glu  
 275 280 285  
 Gly Ser Thr Gly Thr Ser Glu Ser Ser Val Glu Ala Arg Gly Ser Glu  
 290 295 300  
 Glu Val Arg Glu Ser Ser Thr Val Ala Ser Asp Gly Ser Met Glu Gly  
 305 310 315 320  
 Lys Glu Gly Ser Thr Lys Val Glu Glu Asn Ser Met Lys Ala Asp Lys  
 325 330 335  
 Gly Arg Thr Glu Val Asn Gln Cys Ser Ile Asp Leu Gly Glu Asp Asp  
 340 345 350  
 Met Glu Phe Gly Glu Asp Asp Ile Asn Phe Ser Glu Asp Asp Val Glu  
 355 360 365

Ala Val Asn Ile Pro Glu Ser Leu Pro Pro Ser Arg Arg Asn Ser Asn  
 370 375 380

Ser Asn Pro Pro Leu Pro Arg Cys Tyr Gln Cys Lys Ala Ala Lys Val  
 385 390 395 400

Ile Phe Ile Ile Ile Phe Ser Tyr Val Leu Ser Leu Gly Pro Tyr Cys  
 405 410 415

Phe Leu Ala Val Leu Ala Val Trp Val Asp Val Glu Thr Gln Val Pro  
 420 425 430

Gln Trp Val Ile Thr Ile Ile Ile Trp Leu Phe Phe Leu Gln Cys Cys  
 435 440 445

Ile His Pro Tyr Val Tyr Gly Tyr Met His Lys Thr Ile Lys Lys Glu  
 450 455 460

Ile Gln Asp Met Leu Lys Lys Phe Phe Cys Lys Glu Lys Pro Pro Lys  
 465 470 475 480

Glu Asp Ser His Pro Asp Leu Pro Gly Thr Glu Gly Gly Thr Glu Gly  
 485 490 495

Lys Ile Val Pro Ser Tyr Asp Ser Ala Thr Phe Pro  
 500 505

<210> 17  
 <211> 1068  
 <212> DNA  
 <213> Homo sapiens

<400> 17  
 atgcccttga cggacggcat ttcttcattt gaggacctct tggctaacaa taccctcaga 60  
 atatttgtct gggttatagc ttctcattacc tgcttttgaa atctttttgt cattggcatg 120  
 agatctttca ttaaagctga aaatacaact cagctatgt ccatcaaaat cctttgttgc 180  
 gctgattgcc tgatgggtgt ttacttggtc ttgtttggca ttttcgatat aaaataccga 240  
 gggcagtatc agaagtatgc cttgctgtgg atggagagcg tgcaagtccg cctcatgggg 300  
 ttcttgcca tgctgtccac cgaagtctct gtctctgtac tgacctactt gactttggag 360  
 aagttcctgg tcattgtctt ccccttcagt aacattcgac ctggaaaacg gcagacctca 420  
 gtcacacctca ttgcatctg gatggcgga tttttaatag ctgtaattcc attttggaat 480  
 aaggattatt ttggaaactt ttatgggaaa aatggagtat gtttccact ttattatgac 540  
 caaacagaag atattggaag caaagggtat tctcttgga ttttcctagg tgtgaacttg 600  
 ctggcttttc tcatcattgt gttttcctat attactatgt tctgttccat tcaaaaaacc 660  
 gccttgcaga ccacagaagt aaggaattgt ttggaagag aggtggctgt tgcaaatcgt 720  
 ttctttttta tagtggtctc tgatgccatc tgctggattc ctgtatttgt agttaaaatc 780  
 ctttccctct tccgggtgga aataccagac acaatgactt cctggatagt gatttttttc 840  
 cttccagtta acagtgtttt gaatccaatc ctctatactc tcacaaccaa cttttttaag 900  
 gacaagttga aacagctgct gcacaaacat cagaggaaat caattttcaa aattaaaaaa 960

aaaagtttat ctacatccat tgtgtggata gaggactcct cttccctgaa acttgggggtt 1020  
 ttgaacaaaa taacacttgg agacagtata atgaaaccag tttcctag 1068

<210> 18  
 <211> 355  
 <212> PRT  
 <213> Homo sapiens

<400> 18

Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn  
 1 5 10 15  
 Asn Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile Thr Cys Phe  
 20 25 30  
 Gly Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys Ala Glu Asn  
 35 40 45  
 Thr Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu  
 50 55 60  
 Met Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile Lys Tyr Arg  
 65 70 75 80  
 Gly Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Gln Cys  
 85 90 95  
 Arg Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val Ser Val Leu  
 100 105 110  
 Leu Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro  
 115 120 125  
 Phe Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val Ile Leu Ile  
 130 135 140  
 Cys Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro Phe Trp Asn  
 145 150 155 160  
 Lys Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro  
 165 170 175  
 Leu Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly Tyr Ser Leu  
 180 185 190  
 Gly Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe  
 195 200 205  
 Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr  
 210 215 220  
 Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg  
 225 230 235 240  
 Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe  
 245 250 255  
 Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro Asp Thr Met  
 260 265 270  
 Thr Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser Ala Leu Asn



275                      280                      285  
 Pro Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp Lys Leu Lys  
     290                      295                      300  
 Gln Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys Ile Lys Lys  
     305                      310                      315                      320  
 Lys Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser Ser Ser Leu  
                     325                      330                      335  
 Lys Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser Ile Met Lys  
                     340                      345                      350  
 Pro Val Ser  
     355

<210> 19  
 <211> 969  
 <212> DNA  
 <213> Homo sapiens

<400> 19  
 atggatccaa ccattctcaac cttggacaca gaactgacac caatcaacgg aactgaggag 60  
 actctttgct acaagcagac cttgagcctc acggtgctga cgtgcatcgt ttcccttgct 120  
 gggctgacag gaaacgcagt tgtgctctgg ctctgggct gccgcatgcg caggaacgcc 180  
 ttctccatct acatcctcaa cttggccgca gcagacttcc tcttctcag cggccgcctt 240  
 atatattccc tgtaagctt catcagtatc ccccatacca tctctaaaat cctctatcct 300  
 gtgatgatgt ttctctactt tgcaggcctg agctttctga gtgccgtgag caccgagcgc 360  
 tgcctgtccg tcctgtggcc catctggtac cgctgccacc gccccacaca cctgtcagcg 420  
 gtggtgtgtg tcctgctctg ggccctgtcc ctgtgcgga gcatcctgga gtggatgtta 480  
 tgtggcttcc tgttcagtgg tgctgattct gcttggtgtc aaacatcaga ttcatcaca 540  
 gtcgcgtggc tgattttttt atgtgtggtt ctctgtgggt ccagcctggt cctgtgatc 600  
 aggattctct gtggatcccg gaagataccg ctgaccaggc tgtacgtgac catcctgctc 660  
 acagtactgg tcttctctct ctgtggcctg ccctttggca ttcagttttt cctattttta 720  
 tggatccacg tggacaggga agtcttattt tgatcatgtt atctagtttc tattttcctg 780  
 tccgctctta acagcagtgc caaccccatc atttacttct tcgtgggctc ctttaggcag 840  
 cgtcaaaata ggcagaacct gaagctgggt ctccagaggg ctctgcagga cgcgtctgag 900  
 gtggatgaag gtggagggca gcttctctgag gaaatcctgg agctgtcggg aagcagattg 960  
 gagcagtga 969

<210> 20  
 <211> 322  
 <212> PRT  
 <213> Homo sapiens

<400> 20

Met Asp Pro Thr Ile Ser Thr Leu Asp Thr Glu Leu Thr Pro Ile Asn  
 1 5 10 15  
 Gly Thr Glu Glu Thr Leu Cys Tyr Lys Gln Thr Leu Ser Leu Thr Val  
 20 25 30  
 Leu Thr Cys Ile Val Ser Leu Val Gly Leu Thr Gly Asn Ala Val Val  
 35 40 45  
 Leu Trp Leu Leu Gly Cys Arg Met Arg Arg Asn Ala Phe Ser Ile Tyr  
 50 55 60  
 Ile Leu Asn Leu Ala Ala Ala Asp Phe Leu Phe Leu Ser Gly Arg Leu  
 65 70 75 80  
 Ile Tyr Ser Leu Leu Ser Phe Ile Ser Ile Pro His Thr Ile Ser Lys  
 85 90 95  
 Ile Leu Tyr Pro Val Met Met Phe Ser Tyr Phe Ala Gly Leu Ser Phe  
 100 105 110  
 Leu Ser Ala Val Ser Thr Glu Arg Cys Leu Ser Val Leu Trp Pro Ile  
 115 120 125  
 Trp Tyr Arg Cys His Arg Pro Thr His Leu Ser Ala Val Val Cys Val  
 130 135 140  
 Leu Leu Trp Ala Leu Ser Leu Leu Arg Ser Ile Leu Glu Trp Met Leu  
 145 150 155 160  
 Cys Gly Phe Leu Phe Ser Gly Ala Asp Ser Ala Trp Cys Gln Thr Ser  
 165 170 175  
 Asp Phe Ile Thr Val Ala Trp Leu Ile Phe Leu Cys Val Val Leu Cys  
 180 185 190  
 Gly Ser Ser Leu Val Leu Leu Ile Arg Ile Leu Cys Gly Ser Arg Lys  
 195 200 205  
 Ile Pro Leu Thr Arg Leu Tyr Val Thr Ile Leu Leu Thr Val Leu Val  
 210 215 220  
 Phe Leu Leu Cys Gly Leu Pro Phe Gly Ile Gln Phe Phe Leu Phe Leu  
 225 230 235 240  
 Trp Ile His Val Asp Arg Glu Val Leu Phe Cys His Val His Leu Val  
 245 250 255  
 Ser Ile Phe Leu Ser Ala Leu Asn Ser Ser Ala Asn Pro Ile Ile Tyr  
 260 265 270  
 Phe Phe Val Gly Ser Phe Arg Gln Arg Gln Asn Arg Gln Asn Leu Lys  
 275 280 285  
 Leu Val Leu Gln Arg Ala Leu Gln Asp Ala Ser Glu Val Asp Glu Gly  
 290 295 300  
 Gly Gly Gln Leu Pro Glu Glu Ile Leu Glu Leu Ser Gly Ser Arg Leu  
 305 310 315 320  
 Glu Gln

<210> 21  
 <211> 1305

<212> DNA  
 <213> Homo sapiens

<400> 21  
 atggaggatc tcttttagccc ctcaattctg ccgccggcgc ccaacatttc cgtgcccatc 60  
 ttgctgggct ggggtctcaa cctgaccttg gggcaaggag cccctgcctc tgggccgccc 120  
 agccgcgcgc tccgcctggt gttcctgggg gtcatcctgg tggtagcggt ggcaggcaac 180  
 accacagtgc tgtgccgcct gtgcggcggc ggcgggccct ggcggggccc caagcgtcgc 240  
 aagatggact tcctgctggt gcagctggcc ctggcgacc gtacgcgtg cgggggcacg 300  
 gcgctgtcac agctggcctg ggaactgctg ggcgagcccc gcgcggccac gggggacctg 360  
 gcgtgccgct tcctgcagct gctgcaggca tccgggcggg gcgcctcggc ccacctcgtg 420  
 gtgctcatcg ccctcgagcg ccggcgcgcg gtgcgtcttc cgacggccg gccgctgccc 480  
 gcgcgtgccc tcgccgccct gggctggctg ctggcactgc tgcctggcgt gccccggccc 540  
 ttcgtggtgc ggggggactc cccctcgccg ctgccgccgc cgcgcgcc aacgtccctg 600  
 cagccaggcg cgcggggcgc cgcgcgccgc tggccggggg agcgtcgtg ccacgggac 660  
 ttcgcgcccc tgcgcgctg gcacctgcag gtctacgct tctacgaggc cgtcgcgggc 720  
 ttcgtcgcgc ctgttacggt cctgggcgtc gcttgcgccc acctactctc cgtctggtgg 780  
 cggcaccggc cgcaggcccc cgcggctgca gcgcctggt cggcgagccc aggtcgagcc 840  
 cctgcgcccc gcgcgtgccc ccgcgccaag gtgcagagcc tgaagatgag cctgctgctg 900  
 gcgctgctgt tcgtgggctg cgagctgccc tactttgccg cccggctggc ggccgcgtgg 960  
 tcgtccgggc ccgcgggaga ctgggaggga gagggcctgt cggcggcgct gcgcgtggtg 1020  
 gcgatggcca acagcgtct caatcccttc gtctacctt tcttcaggc gggcgactgc 1080  
 cggtccggc gacagctgc gaagcggctg ggctctctgt gctgcgcgcc gcagggaggc 1140  
 gcggaggacg aggaggggcc ccggggccac caggcgtct accgccaacg ctggccccac 1200  
 cctcattatc accatgctcg gcgggaaccg ctggacgagg gcggttgcg cccaccccct 1260  
 ccgcgccccca gaccctgcc ttgctcctgc gaaagtgcct tctag 1305

<210> 22  
 <211> 434  
 <212> PRT  
 <213> Homo sapiens

<400> 22

Met Glu Asp Leu Phe Ser Pro Ser Ile Leu Pro Pro Ala Pro Asn Ile  
 1 5 10 15  
 Ser Val Pro Ile Leu Leu Gly Trp Gly Leu Asn Leu Thr Leu Gly Gln  
 20 25 30  
 Gly Ala Pro Ala Ser Gly Pro Pro Ser Arg Arg Val Arg Leu Val Phe  
 35 40 45

Leu Gly Val Ile Leu Val Val Ala Val Ala Gly Asn Thr Thr Val Leu  
 50 55 60  
 Cys Arg Leu Cys Gly Gly Gly Gly Pro Trp Ala Gly Pro Lys Arg Arg  
 65 70 75 80  
 Lys Met Asp Phe Leu Leu Val Gln Leu Ala Leu Ala Asp Leu Tyr Ala  
 85 90 95  
 Cys Gly Gly Thr Ala Leu Ser Gln Leu Ala Trp Glu Leu Leu Gly Glu  
 100 105 110  
 Pro Arg Ala Ala Thr Gly Asp Leu Ala Cys Arg Phe Leu Gln Leu Leu  
 115 120 125  
 Gln Ala Ser Gly Arg Gly Ala Ser Ala His Leu Val Val Leu Ile Ala  
 130 135 140  
 Leu Glu Arg Arg Arg Ala Val Arg Leu Pro His Gly Arg Pro Leu Pro  
 145 150 155 160  
 Ala Arg Ala Leu Ala Ala Leu Gly Trp Leu Leu Ala Leu Leu Leu Ala  
 165 170 175  
 Leu Pro Pro Ala Phe Val Val Arg Gly Asp Ser Pro Ser Pro Leu Pro  
 180 185 190  
 Pro Pro Pro Pro Pro Thr Ser Leu Gln Pro Gly Ala Pro Pro Ala Ala  
 195 200 205  
 Arg Ala Trp Pro Gly Glu Arg Arg Cys His Gly Ile Phe Ala Pro Leu  
 210 215 220  
 Pro Arg Trp His Leu Gln Val Tyr Ala Phe Tyr Glu Ala Val Ala Gly  
 225 230 235 240  
 Phe Val Ala Pro Val Thr Val Leu Gly Val Ala Cys Gly His Leu Leu  
 245 250 255  
 Ser Val Trp Trp Arg His Arg Pro Gln Ala Pro Ala Ala Ala Pro  
 260 265 270  
 Trp Ser Ala Ser Pro Gly Arg Ala Pro Ala Pro Ser Ala Leu Pro Arg  
 275 280 285  
 Ala Lys Val Gln Ser Leu Lys Met Ser Leu Leu Leu Ala Leu Leu Phe  
 290 295 300  
 Val Gly Cys Glu Leu Pro Tyr Phe Ala Ala Arg Leu Ala Ala Ala Trp  
 305 310 315 320  
 Ser Ser Gly Pro Ala Gly Asp Trp Glu Gly Glu Gly Leu Ser Ala Ala  
 325 330 335  
 Leu Arg Val Val Ala Met Ala Asn Ser Ala Leu Asn Pro Phe Val Tyr  
 340 345 350  
 Leu Phe Phe Gln Ala Gly Asp Cys Arg Leu Arg Arg Gln Leu Arg Lys  
 355 360 365  
 Arg Leu Gly Ser Leu Cys Cys Ala Pro Gln Gly Gly Ala Glu Asp Glu  
 370 375 380  
 Glu Gly Pro Arg Gly His Gln Ala Leu Tyr Arg Gln Arg Trp Pro His  
 385 390 395 400

Pro His Tyr His His Ala Arg Arg Glu Pro Leu Asp Glu Gly Gly Leu  
 405 410 415

Arg Pro Pro Pro Arg Pro Arg Pro Leu Pro Cys Ser Cys Glu Ser  
 420 425 430

Ala Phe

<210> 23  
 <211> 1041  
 <212> DNA  
 <213> Homo sapiens

<400> 23  
 atgtacaacg ggtcgtgctg ccgcatcgag ggggacacca tctcccaggt gatgccgccg 60  
 ctgctcattg tggcctttgt gctgggcgca ctaggcaatg gggtcgccct gtgtggtttc 120  
 tgcttcacaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct 180  
 gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac 240  
 tgggcttttg gggacattcc ctgccgagtg gggctcttca cggtggccat gaacagggcc 300  
 gggagcatcg tgctccttac ggtggtggct gcggacaggt atttcaaagt ggtccacccc 360  
 caccacgcgg tgaacactat ctccaccggt gtggcggtcg gcatcgctcg caccctgtgg 420  
 gccctggtca tcttggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag 480  
 acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg 540  
 ttccagctgg agttctttat gccctcggc atcatcttat tttgctcctt caagattggt 600  
 tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgaccggt 660  
 ttcatcatgg tggtggcaat tgtgttcac acatgctacc tgcccagcgt gtctgctaga 720  
 ctctatttcc tctggacggt gccctcgagt gcctgcgac cctctgtcca tggggccctg 780  
 cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt 840  
 tcaagccct cctttcccaa attctacaac aagctcaaaa tctgcagtct gaaacccaag 900  
 cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcgaa cctcgggtcgc 960  
 aggagtgc tcaagtgtggc aaatagtctt caaagccagt ctgatgggca atgggatccc 1020  
 cacattgttg agtggcactg a 1041

<210> 24  
 <211> 346  
 <212> PRT  
 <213> Homo sapiens

<400> 24

Met Tyr Asn Gly Ser Cys Cys Arg Ile Glu Gly Asp Thr Ile Ser Gln  
 1 5 10 15

Val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly  
 20 25 30

Asn Gly Val Ala Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys  
 35 40 45  
 Pro Ser Thr Val Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu  
 50 55 60  
 Met Ile Cys Leu Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His  
 65 70 75 80  
 Trp Ala Phe Gly Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala  
 85 90 95  
 Met Asn Arg Ala Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp  
 100 105 110  
 Arg Tyr Phe Lys Val Val His Pro His His Ala Val Asn Thr Ile Ser  
 115 120 125  
 Thr Arg Val Ala Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile  
 130 135 140  
 Leu Gly Thr Val Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu  
 145 150 155 160  
 Thr Ala Val Ser Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp  
 165 170 175  
 His Asp Ile Met Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile  
 180 185 190  
 Leu Phe Cys Ser Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln  
 195 200 205  
 Leu Ala Arg Gln Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val  
 210 215 220  
 Val Ala Ile Val Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg  
 225 230 235 240  
 Leu Tyr Phe Leu Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val  
 245 250 255  
 His Gly Ala Leu His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met  
 260 265 270  
 Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe  
 275 280 285  
 Tyr Asn Lys Leu Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His  
 290 295 300  
 Ser Lys Thr Gln Arg Pro Glu Glu Met Pro Ile Ser Asn Leu Gly Arg  
 305 310 315 320  
 Arg Ser Cys Ile Ser Val Ala Asn Ser Phe Gln Ser Gln Ser Asp Gly  
 325 330 335  
 Gln Trp Asp Pro His Ile Val Glu Trp His  
 340 345

<210> 25  
 <211> 1011  
 <212> DNA  
 <213> Homo sapiens

<400> 25  
 atgaacaaca atacaacatg tattcaacca tctatgatct cttccatggc tttaaccaatc 60  
 atttacatcc tcctttgtat tgttggtgtt tttggaaca ctctctctca atggatattt 120  
 ttaacaaaaa taggtaaaaa aacatcaacg cacatctacc tgtcacacct tgtgactgca 180  
 aacttacttg tgtgcagtgc catgcctttc atgagtatct atttcctgaa aggtttccaa 240  
 tgggaatatc aatctgctca atgcagagtg gtcaattttc tgggaactct atccatgcat 300  
 gcaagtatgt ttgtcagtct ctttaatttta agttggattg ccataagccg ctatgctacc 360  
 ttaatgcaaa aggattcctc gcaagagact acttcatgct atgagaaaaat attttatggc 420  
 catttactga aaaaatttcg ccagcccaac tttgctagaa aactatgcat ttacatatgg 480  
 ggagttgtac tgggcataat cattccagtt accgtatact actcagtcac agaggctaca 540  
 gaaggagaag agagcctatg ctacaatcgg cagatggaac taggagccat gatctctcag 600  
 attgcaggtc tcattggaac cacatttatt ggattttcct ttttagtagt actaacatca 660  
 tactactctt ttgtaagcca tctgagaaaa ataagaacct gtacgtccat tatggagaaa 720  
 gatttgactt acagttctgt gaaaagacat cttttggtca tccagattct actaatagtt 780  
 tgcttccttc cttatagtat ttttaaacc cttttttatg ttctacacca aagagataac 840  
 tgtcagcaat tgaattattt aatagaaaca aaaacattc tcacctgtct tgcttcggcc 900  
 agaagtagca cagaccccat tatatttctt ttattagata aaacattcaa gaagacacta 960  
 tataatctct ttacaaagtc taattcagca catatgcaat catatggttg a 1011

<210> 26  
 <211> 336  
 <212> PRT  
 <213> Homo sapiens

<400> 26  
 Met Asn Asn Asn Thr Thr Cys Ile Gln Pro Ser Met Ile Ser Ser Met  
 1 5 10 15  
 Ala Leu Pro Ile Ile Tyr Ile Leu Leu Cys Ile Val Gly Val Phe Gly  
 20 25 30  
 Asn Thr Leu Ser Gln Trp Ile Phe Leu Thr Lys Ile Gly Lys Lys Thr  
 35 40 45  
 Ser Thr His Ile Tyr Leu Ser His Leu Val Thr Ala Asn Leu Leu Val  
 50 55 60  
 Cys Ser Ala Met Pro Phe Met Ser Ile Tyr Phe Leu Lys Gly Phe Gln  
 65 70 75 80  
 Trp Glu Tyr Gln Ser Ala Gln Cys Arg Val Val Asn Phe Leu Gly Thr  
 85 90 95  
 Leu Ser Met His Ala Ser Met Phe Val Ser Leu Leu Ile Leu Ser Trp  
 100 105 110

Ile Ala Ile Ser Arg Tyr Ala Thr Leu Met Gln Lys Asp Ser Ser Gln  
 115 120 125  
 Glu Thr Thr Ser Cys Tyr Glu Lys Ile Phe Tyr Gly His Leu Leu Lys  
 130 135 140  
 Lys Phe Arg Gln Pro Asn Phe Ala Arg Lys Leu Cys Ile Tyr Ile Trp  
 145 150 155 160  
 Gly Val Val Leu Gly Ile Ile Ile Pro Val Thr Val Tyr Tyr Ser Val  
 165 170 175  
 Ile Glu Ala Thr Glu Gly Glu Glu Ser Leu Cys Tyr Asn Arg Gln Met  
 180 185 190  
 Glu Leu Gly Ala Met Ile Ser Gln Ile Ala Gly Leu Ile Gly Thr Thr  
 195 200 205  
 Phe Ile Gly Phe Ser Phe Leu Val Val Leu Thr Ser Tyr Tyr Ser Phe  
 210 215 220  
 Val Ser His Leu Arg Lys Ile Arg Thr Cys Thr Ser Ile Met Glu Lys  
 225 230 235 240  
 Asp Leu Thr Tyr Ser Ser Val Lys Arg His Leu Leu Val Ile Gln Ile  
 245 250 255  
 Leu Leu Ile Val Cys Phe Leu Pro Tyr Ser Ile Phe Lys Pro Ile Phe  
 260 265 270  
 Tyr Val Leu His Gln Arg Asp Asn Cys Gln Gln Leu Asn Tyr Leu Ile  
 275 280 285  
 Glu Thr Lys Asn Ile Leu Thr Cys Leu Ala Ser Ala Arg Ser Ser Thr  
 290 295 300  
 Asp Pro Ile Ile Phe Leu Leu Leu Asp Lys Thr Phe Lys Lys Thr Leu  
 305 310 315 320  
 Tyr Asn Leu Phe Thr Lys Ser Asn Ser Ala His Met Gln Ser Tyr Gly  
 325 330 335

<210> 27  
 <211> 1014  
 <212> DNA  
 <213> Homo sapiens

<400> 27  
 atgaatgagc cactagacta tttagcaaat gcttctgatt tccccgatta tgcagctgct 60  
 tttggaaatt gcactgatga aaacatccca ctcaagatgc actacctccc tggtatttat 120  
 ggcattatct tcctcgtggg atttccaggc aatgcagtag tgatatccac ttacattttc 180  
 aaaatgagac cttggaagag cagcaccatc attatgctga acctggcctg cacagatctg 240  
 ctgtatctga ccagcctccc ctctctgatt cactactatg ccagtggcga aaactggatc 300  
 tttggagatt tcattgtgtaa gtttatccgc ttcagcttcc atttcaacct gtatagcagc 360  
 atcctcttcc tcacctgttt cagcatcttc cgctactgtg tgatcattca cccaatgagc 420  
 tgcttttcca ttcacaaaac tcgatgtgca gttgtagcct gtgctgtggt gtggatcatt 480  
 tcaactgtag ctgtcattcc gatgaccttc ttgatcacat caaccaacag gaccaacaga 540



```

tcagcctgtc tcgacctcac cagttcggat gaactcaata ctattaagtg gtacaacctg 600
attttgactg caactacttt ctgcctcccc ttggtgatag tgacactttg ctataccacg 660
attatccaca ctctgaccca tggactgcaa actgacagct gccttaagca gaaagcacga 720
aggctaacca ttctgctact ccttgcattht tacgtatgtht ttttaccctt ccatatcttg 780
agggtcattc ggatcgaatc tcgcctgctt tcaatcagtht gttccattga gaatcagatc 840
catgaagctt acatcgthtt tagaccatta gctgctctga acacctttgg taacctgtta 900
ctatatgtgg tggtcagcga caactttcag caggctgtct gctcaacagt gagatgcaaa 960
gtaagcggga accttgagca agcaaagaaa attagttact caaacaaccc ttga 1014

```

```

<210> 28
<211> 337
<212> PRT
<213> Homo sapiens

```

```

<400> 28

```

```

Met Asn Glu Pro Leu Asp Tyr Leu Ala Asn Ala Ser Asp Phe Pro Asp
1          5          10          15
Tyr Ala Ala Ala Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro Leu Lys
          20          25          30
Met His Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val Gly Phe
          35          40          45
Pro Gly Asn Ala Val Val Ile Ser Thr Tyr Ile Phe Lys Met Arg Pro
          50          55          60
Trp Lys Ser Ser Thr Ile Ile Met Leu Asn Leu Ala Cys Thr Asp Leu
65          70          75          80
Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala Ser Gly
          85          90          95
Glu Asn Trp Ile Phe Gly Asp Phe Met Cys Lys Phe Ile Arg Phe Ser
          100          105          110
Phe His Phe Asn Leu Tyr Ser Ser Ile Leu Phe Leu Thr Cys Phe Ser
          115          120          125
Ile Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile
          130          135          140
His Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile
145          150          155          160
Ser Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn
          165          170          175
Arg Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu
          180          185          190
Asn Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Thr Thr Phe Cys
          195          200          205
Leu Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr

```

210                      215                      220  
 Leu Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg  
 225                      230                      235                      240  
 Arg Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro  
 245                      250                      255  
 Phe His Ile Leu Arg Val Ile Arg Ile Glu Ser Arg Leu Leu Ser Ile  
 260                      265                      270  
 Ser Cys Ser Ile Glu Asn Gln Ile His Glu Ala Tyr Ile Val Ser Arg  
 275                      280                      285  
 Pro Leu Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val Val  
 290                      295                      300  
 Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys Lys  
 305                      310                      315                      320  
 Val Ser Gly Asn Leu Glu Gln Ala Lys Lys Ile Ser Tyr Ser Asn Asn  
 325                      330                      335

Pro

<210> 29  
 <211> 993  
 <212> DNA  
 <213> Homo sapiens

<400> 29  
 atggatccaa ccaccccggc ctggggaaca gaaagtacaa cagtgaatgg aaatgaccaa 60  
 gcccttcttc tgctttgtgg caaggagacc ctgatcccg tcttctgat ccttttcatt 120  
 gccctggctg ggctggtagg aaacgggttt gtgctctggc tcctgggctt ccgcatgcgc 180  
 aggaacgcct tctctgtcta cgtcctcagc ctggccgggg ccgacttcct ctctctctgc 240  
 ttccagatta taaattgcct ggtgtacctc agtaacttct tctgttccat ctccatcaat 300  
 ttccctagct tcttcaccac tgtgatgacc tgtgcctacc ttgcaggcct gagcatgctg 360  
 agcaccgtca gcaccgagcg ctgcctgtcc gtccctgtgg ccactctggt tcgctgccgc 420  
 cgccccagac acctgtcagc ggtcgtgtgt gtccctgtct gggccctgtc cctactgctg 480  
 agcatcttgg aagggaaagt ctgtggcttc ttatttagtg atggtgactc tggttggtgt 540  
 cagacatttg atttcatcac tgcagcgtgg ctgatttttt tattcatggt tctctgtggg 600  
 tccagtctgg ccctgctggt caggatcctc tgtggctcca ggggtctgcc actgaccagg 660  
 ctgtacctga ccatcctgct cacagtgctg gtgttctctc tctgcggcct gccctttggc 720  
 attcagtggg tcctaattatt atggatctgg aaggattctg atgtcttatt ttgtcatatt 780  
 catccagttt cagttgtcct gtcattctct aacagcagtg ccaaccccat catttacttc 840  
 ttcgtgggct cttttaggaa gcagtggcgg ctgcagcagc cgatcctcaa gctggctctc 900  
 cagagggctc tgcaggacat tgctgaggtg gatcacagtg aaggatgctt ccgtcagggc 960  
 accccggaga tgtcgagaag cagtctggtg tag 993

<210> 30  
 <211> 330  
 <212> PRT  
 <213> Homo sapiens

<400> 30

```

Met Asp Pro Thr Thr Pro Ala Trp Gly Thr Glu Ser Thr Thr Val Asn
1          5          10          15

Gly Asn Asp Gln Ala Leu Leu Leu Leu Cys Gly Lys Glu Thr Leu Ile
          20          25          30

Pro Val Phe Leu Ile Leu Phe Ile Ala Leu Val Gly Leu Val Gly Asn
          35          40          45

Gly Phe Val Leu Trp Leu Leu Gly Phe Arg Met Arg Arg Asn Ala Phe
          50          55          60

Ser Val Tyr Val Leu Ser Leu Ala Gly Ala Asp Phe Leu Phe Leu Cys
65          70          75          80

Phe Gln Ile Ile Asn Cys Leu Val Tyr Leu Ser Asn Phe Phe Cys Ser
          85          90          95

Ile Ser Ile Asn Phe Pro Ser Phe Phe Thr Thr Val Met Thr Cys Ala
          100          105          110

Tyr Leu Ala Gly Leu Ser Met Leu Ser Thr Val Ser Thr Glu Arg Cys
          115          120          125

Leu Ser Val Leu Trp Pro Ile Trp Tyr Arg Cys Arg Arg Pro Arg His
          130          135          140

Leu Ser Ala Val Val Cys Val Leu Leu Trp Ala Leu Ser Leu Leu Leu
145          150          155          160

Ser Ile Leu Glu Gly Lys Phe Cys Gly Phe Leu Phe Ser Asp Gly Asp
          165          170          175

Ser Gly Trp Cys Gln Thr Phe Asp Phe Ile Thr Ala Ala Trp Leu Ile
          180          185          190

Phe Leu Phe Met Val Leu Cys Gly Ser Ser Leu Ala Leu Leu Val Arg
          195          200          205

Ile Leu Cys Gly Ser Arg Gly Leu Pro Leu Thr Arg Leu Tyr Leu Thr
210          215          220

Ile Leu Leu Thr Val Leu Val Phe Leu Leu Cys Gly Leu Pro Phe Gly
225          230          235          240

Ile Gln Trp Phe Leu Ile Leu Trp Ile Trp Lys Asp Ser Asp Val Leu
          245          250          255

Phe Cys His Ile His Pro Val Ser Val Val Leu Ser Ser Leu Asn Ser
          260          265          270

Ser Ala Asn Pro Ile Ile Tyr Phe Phe Val Gly Ser Phe Arg Lys Gln
          275          280          285

Trp Arg Leu Gln Gln Pro Ile Leu Lys Leu Ala Leu Gln Arg Ala Leu
290          295          300

```

Gln Asp Ile Ala Glu Val Asp His Ser Glu Gly Cys Phe Arg Gln Gly  
305 310 315 320

Thr Pro Glu Met Ser Arg Ser Ser Leu Val  
325 330

<210> 31  
<211> 1092  
<212> DNA  
<213> Homo sapiens

<400> 31  
atgggccccg gcgaggcgct gctggcgggt ctccctggtga tgggtactggc cgtggcgctg 60  
ctatccaacg cactgggtgct gctttgttgc gcctacagcg ctgagctccg cactcgagcc 120  
tcaggcgtcc tcctgggtgaa tctgtcgctg ggccacctgc tgctggcggc gctggacatg 180  
cccttcacgc tgctcgggtg gatgcgcggg cggacaccgt cggcgcccg cgcatgccaa 240  
gtcattggct tcctggacac cttcctggcg tccaacgcgg cgctgagcgt ggcggcgctg 300  
agcgagacc agtggctggc agtgggcttc ccactgcgct acgccggacg cctgcgaccg 360  
cgctatgccg gcctgctgct gggctgtgcc tggggacagt cgctggcctt ctcaggcgct 420  
gcacttggct gctcgtggct tggctacagc agcgccttcg cgtcctgttc gctgcgcctg 480  
ccgcccagc ctgagcgtcc gcgcttcgca gccttcaccg ccacgctcca tgccgtgggc 540  
ttcgtgctgc cgctggcggt gctctgcctc acctcgctcc aggtgcaccg ggtggcacgc 600  
agccactgcc agcgcatgga caccgtcacc atgaaggcgc tcgcgctgct cgccgacctg 660  
caccaccagt tgccgcagcg ctgcctcacc cagcagaagc ggcgccgcca ccgcgccacc 720  
aggaagattg gcattgctat tgcgacctc ctcactgtgt ttgccccgta tgcatgacc 780  
aggtggcgg agctcgtgcc cttcgtcacc gtgaacgccc agtggggcat cctcagcaag 840  
tgctgacct acagcaaggc ggtggccgac ccgttcacgt actctctgct ccgccggccg 900  
ttccgccaag tcctggcccg catggtgcac cggtgtgta agagaacccc gcgccagca 960  
tccaccatg acagctctct ggatgtggc ggcatggtgc accagctgct gaagagaacc 1020  
ccgcgccag cgtccacca caacggctct gtggacacag agaatgattc ctgcctgcag 1080  
cagacacact ga 1092

<210> 32  
<211> 363  
<212> PRT  
<213> Homo sapiens

<400> 32

Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met Val Leu  
1 5 10 15  
Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr  
20 25 30

Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu  
 35 40 45  
 Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu  
 50 55 60  
 Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln  
 65 70 75 80  
 Val Ile Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala Leu Ser  
 85 90 95  
 Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe Pro Leu  
 100 105 110  
 Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu Leu Gly  
 115 120 125  
 Cys Ala Trp Gly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leu Gly Cys  
 130 135 140  
 Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu Arg Leu  
 145 150 155 160  
 Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala Thr Leu  
 165 170 175  
 His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu Thr Ser  
 180 185 190  
 Leu Gln Val His Arg Val Ala Arg Ser His Cys Gln Arg Met Asp Thr  
 195 200 205  
 Val Thr Met Lys Ala Leu Ala Leu Leu Ala Asp Leu His Pro Ser Val  
 210 215 220  
 Arg Gln Arg Cys Leu Ile Gln Gln Lys Arg Arg Arg His Arg Ala Thr  
 225 230 235 240  
 Arg Lys Ile Gly Ile Ala Ile Ala Thr Phe Leu Ile Cys Phe Ala Pro  
 245 250 255  
 Tyr Val Met Thr Arg Leu Ala Glu Leu Val Pro Phe Val Thr Val Asn  
 260 265 270  
 Ala Gln Trp Gly Ile Leu Ser Lys Cys Leu Thr Tyr Ser Lys Ala Val  
 275 280 285  
 Ala Asp Pro Phe Thr Tyr Ser Leu Leu Arg Arg Pro Phe Arg Gln Val  
 290 295 300  
 Leu Ala Gly Met Val His Arg Leu Leu Lys Arg Thr Pro Arg Pro Ala  
 305 310 315 320  
 Ser Thr His Asp Ser Ser Leu Asp Val Ala Gly Met Val His Gln Leu  
 325 330 335  
 Leu Lys Arg Thr Pro Arg Pro Ala Ser Thr His Asn Gly Ser Val Asp  
 340 345 350  
 Thr Glu Asn Asp Ser Cys Leu Gln Gln Thr His  
 355 360

<210> 33  
 <211> 1125

<212> DNA  
<213> Homo sapiens

<400> 33  
atgccacac tcaatacttc tgcctctcca cccacattct tctgggcca tgcctccgga 60  
ggcagtgctg tgagtgtgta tgatgtctcg atgcctgtca aattcctagc cctgaggctc 120  
atggttgccc tggcctatgg gcttgtgggg gccattggct tgctgggaaa tttggcggtg 180  
ctgtgggtac tgagtaactg tgcccggaga gcccctggcc caccttcaga caccttcgtc 240  
ttcaacctgg ctctggcgga cctgggactg gcactcactc tccccttttg ggagccgag 300  
tcggcactgg actttcactg gcccttcgga ggtgccctct gcaagatggt tctgacggcc 360  
actgtcctca acgtctatgc cagcatcttc ctcatcacag cgctgagcgt tgctcgctac 420  
tggtgtgtgg ccatggctgc ggggccaggc acccacctct cactcttctg ggccgaata 480  
gccaccctgg cagtgtgggc ggcggctgcc ctggtgacgg tgcccacagc tgtcttcggg 540  
gtggaggggt aggtgtgtgg tgtgcgcctt tgcctgtctg gtttccccag caggtagctg 600  
ctgggggcct accagctgca gaggggtgtg ctggctttca tgggtgccctt gggcgtcatc 660  
accaccagct acctgtgtgt gctggccttc ctgcagcggc ggcaacggcg gcggcaggac 720  
agcagggctg tggcccgtc tgtccgcac ctggtggctt ccttcttcct ctgctgggtt 780  
cccaaccatg tggcactct ctgggtgtgc ctggtgaagt ttgacctggt gccctggaac 840  
agtactttct atactatcca gacgtatgtc ttccctgtca ctacttgctt ggcacacagc 900  
aatagctgcc tcaaccctgt gctgtactgt ctccctgagg gggagccccg gcaggctctg 960  
gcaggcacct tcagggatct gcggtcgagg ctgtggcccc agggcgagg ctgggtgcaa 1020  
caggtggccc taaagcaggt aggcaggcgg tgggtcgcaa gcaacccccg ggagagccgc 1080  
ccttctaccc tgctcaccaa cctggacaga gggacaccgc ggtga 1125

<210> 34  
<211> 374  
<212> PRT  
<213> Homo sapiens

<400> 34  
Met Pro Thr Leu Asn Thr Ser Ala Ser Pro Pro Thr Phe Phe Trp Ala  
1 5 10 15  
Asn Ala Ser Gly Gly Ser Val Leu Ser Ala Asp Asp Ala Pro Met Pro  
20 25 30  
Val Lys Phe Leu Ala Leu Arg Leu Met Val Ala Leu Ala Tyr Gly Leu  
35 40 45  
Val Gly Ala Ile Gly Leu Leu Gly Asn Leu Ala Val Leu Trp Val Leu  
50 55 60  
Ser Asn Cys Ala Arg Arg Ala Pro Gly Pro Pro Ser Asp Thr Phe Val  
65 70 75 80

Phe Asn Leu Ala Leu Ala Asp Leu Gly Leu Ala Leu Thr Leu Pro Phe  
                     85                    90                    95  
 Trp Ala Ala Glu Ser Ala Leu Asp Phe His Trp Pro Phe Gly Gly Ala  
                     100                    105                    110  
 Leu Cys Lys Met Val Leu Thr Ala Thr Val Leu Asn Val Tyr Ala Ser  
                     115                    120                    125  
 Ile Phe Leu Ile Thr Ala Leu Ser Val Ala Arg Tyr Trp Val Val Ala  
                     130                    135                    140  
 Met Ala Ala Gly Pro Gly Thr His Leu Ser Leu Phe Trp Ala Arg Ile  
                     145                    150                    155                    160  
 Ala Thr Leu Ala Val Trp Ala Ala Ala Ala Leu Val Thr Val Pro Thr  
                     165                    170                    175  
 Ala Val Phe Gly Val Glu Gly Glu Val Cys Gly Val Arg Leu Cys Leu  
                     180                    185                    190  
 Leu Arg Phe Pro Ser Arg Tyr Trp Leu Gly Ala Tyr Gln Leu Gln Arg  
                     195                    200                    205  
 Val Val Leu Ala Phe Met Val Pro Leu Gly Val Ile Thr Thr Ser Tyr  
                     210                    215                    220  
 Leu Leu Leu Leu Ala Phe Leu Gln Arg Arg Gln Arg Arg Arg Gln Asp  
                     225                    230                    235                    240  
 Ser Arg Val Val Ala Arg Ser Val Arg Ile Leu Val Ala Ser Phe Phe  
                     245                    250                    255  
 Leu Cys Trp Phe Pro Asn His Val Val Thr Leu Trp Gly Val Leu Val  
                     260                    265                    270  
 Lys Phe Asp Leu Val Pro Trp Asn Ser Thr Phe Tyr Thr Ile Gln Thr  
                     275                    280                    285  
 Tyr Val Phe Pro Val Thr Thr Cys Leu Ala His Ser Asn Ser Cys Leu  
                     290                    295                    300  
 Asn Pro Val Leu Tyr Cys Leu Leu Arg Arg Glu Pro Arg Gln Ala Leu  
                     305                    310                    315                    320  
 Ala Gly Thr Phe Arg Asp Leu Arg Ser Arg Leu Trp Pro Gln Gly Gly  
                     325                    330                    335  
 Gly Trp Val Gln Gln Val Ala Leu Lys Gln Val Gly Arg Arg Trp Val  
                     340                    345                    350  
 Ala Ser Asn Pro Arg Glu Ser Arg Pro Ser Thr Leu Leu Thr Asn Leu  
                     355                    360                    365  
 Asp Arg Gly Thr Pro Gly  
                     370  
 <210> 35  
 <211> 1092  
 <212> DNA  
 <213> Homo sapiens  
 <400> 35  
 atgaatcggc accatctgca ggatcacttt ctggaaatag acaagaagaa ctgctgtgtg

60

```

ttccgagatg acttcattgt caaggtgttg ccgccggtgt tggggctgga gtttatcttc 120
gggcttcttg gcaatggcct tgccctgtgg attttctgtt tccacctcaa gtcctggaaa 180
tccagccgga ttttctgtt caacctggca gtggctgact ttctactgat catctgcctg 240
cccttcctga tggacaacta tgtgaggcgt tgggactgga agtttgggga catcccttgc 300
cggctgatgc tcttcattgt ggctatgaac cgccagggca gcatcatctt cctcacggtg 360
gtggcggtag acaggtatgt ccgggtgttc catccccacc acgccctgaa caagatctcc 420
aatcggacag cagccatcat ctcttgctt ctgtggggca tcactattgg cctgacagtc 480
cacctcctga agaagaagat gccgatccag aatggcggtg caaatttgtg cagcagcttc 540
agcatctgcc ataccttcca gtggcacgaa gccatgttcc tcctggagtt ctctctgccc 600
ctgggcatca tcctgttctg ctccagccaga attatctgga gcctgcggca gagacaaatg 660
gaccggcatg ccaagatcaa gagagccatc acctcatca tgggtgtggc catcgtcttt 720
gtcatctgct tccttcccag cgtggttgtg cggatccgca tcttctggct cctgcacact 780
tcgggcacgc agaattgtga agtgtaccgc tcgggtggacc tggcggtctt tatcactctc 840
agcttcacct acatgaacag catgctggac ccgctggtgt actacttctc cagcccatcc 900
tttccaact tcttctccac tttgatcaac cgctgcctcc agaggaagat gacaggtgag 960
ccagataata accgcagcac gagcgtcgag ctccagggg accccaacaa aaccagaggc 1020
gctccagagg cgtaaatggc caactccggt gagccatgga gccctctta tctgggcccc 1080
acctctcctt aa 1092

```

<210> 36  
 <211> 363  
 <212> PRT  
 <213> Homo sapiens

<400> 36

```

Met Asn Arg His His Leu Gln Asp His Phe Leu Glu Ile Asp Lys Lys
1          5          10          15
Asn Cys Cys Val Phe Arg Asp Asp Phe Ile Val Lys Val Leu Pro Pro
20          25          30
Val Leu Gly Leu Glu Phe Ile Phe Gly Leu Leu Gly Asn Gly Leu Ala
35          40          45
Leu Trp Ile Phe Cys Phe His Leu Lys Ser Trp Lys Ser Ser Arg Ile
50          55          60
Phe Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu Ile Ile Cys Leu
65          70          75          80
Pro Phe Leu Met Asp Asn Tyr Val Arg Arg Trp Asp Trp Lys Phe Gly
85          90          95
Asp Ile Pro Cys Arg Leu Met Leu Phe Met Leu Ala Met Asn Arg Gln
100          105          110

```



Gly Ser Ile Ile Phe Leu Thr Val Val Ala Val Asp Arg Tyr Phe Arg  
           115                                  120                                  125  
 Val Val His Pro His His Ala Leu Asn Lys Ile Ser Asn Arg Thr Ala  
           130                                  135                                  140  
 Ala Ile Ile Ser Cys Leu Leu Trp Gly Ile Thr Ile Gly Leu Thr Val  
           145                                  150                                  155                                  160  
 His Leu Leu Lys Lys Lys Met Pro Ile Gln Asn Gly Gly Ala Asn Leu  
                                   165                                  170                                  175  
 Cys Ser Ser Phe Ser Ile Cys His Thr Phe Gln Trp His Glu Ala Met  
                                   180                                  185                                  190  
 Phe Leu Leu Glu Phe Phe Leu Pro Leu Gly Ile Ile Leu Phe Cys Ser  
                                   195                                  200                                  205  
 Ala Arg Ile Ile Trp Ser Leu Arg Gln Arg Gln Met Asp Arg His Ala  
           210                                  215                                  220  
 Lys Ile Lys Arg Ala Ile Thr Phe Ile Met Val Val Ala Ile Val Phe  
           225                                  230                                  235                                  240  
 Val Ile Cys Phe Leu Pro Ser Val Val Val Arg Ile Arg Ile Phe Trp  
                                   245                                  250                                  255  
 Leu Leu His Thr Ser Gly Thr Gln Asn Cys Glu Val Tyr Arg Ser Val  
                                   260                                  265                                  270  
 Asp Leu Ala Phe Phe Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met  
           275                                  280                                  285  
 Leu Asp Pro Val Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Asn Phe  
           290                                  295                                  300  
 Phe Ser Thr Leu Ile Asn Arg Cys Leu Gln Arg Lys Met Thr Gly Glu  
           305                                  310                                  315                                  320  
 Pro Asp Asn Asn Arg Ser Thr Ser Val Glu Leu Thr Gly Asp Pro Asn  
                                   325                                  330                                  335  
 Lys Thr Arg Gly Ala Pro Glu Ala Leu Met Ala Asn Ser Gly Glu Pro  
                                   340                                  345                                  350  
 Trp Ser Pro Ser Tyr Leu Gly Pro Thr Ser Pro  
           355                                  360

<210> 37  
 <211> 1044  
 <212> DNA  
 <213> Homo sapiens

<400> 37  
 atgggggatg agctggcacc ttgccctgtg ggcactacag cttggccggc cctgatccag 60  
 ctcacacagca agacaccctg catgccccaa gcagccagca acacttcctt gggcctgggg 120  
 gacctcaggg tgcccagctc catgctgtac tggcttttcc ttccctcaag cctgctggct 180  
 gcagccacac tggctgtcag cccctgtctg ctgggtgacca tcctgcggaa ccaacggctg 240  
 cgacaggagc cccactacct gctcccggt aacatcctgc tctcagacct ggcctacatt 300  
 ctctccaca tgctcatctc ctccagcagc ctgggtgggt gggagctggg ccgcatggcc 360

```

tgtggcattc tcaactgatgc tgtcttcgcc gcctgcacca gcaccatcct gtccttcacc 420
gccattgtgc tgcacaccta cctggcagtc atccatccac tgcgctacct ctccttcattg 480
tcccatgggg ctgcctggaa ggcagtggcc ctcatctggc tgggtggcctg ctgcttcccc 540
acattcctta tttggctcag caagtggcag gatgcccagc tggaggagca aggagcttca 600
tacatcctac caccaagcat gggcaccag ccgggatgtg gcctcctggt cattgttacc 660
tacacctcca ttctgtgcgt tctgttcctc tgcacagctc tcattgcca ctgtttctgg 720
aggatctatg cagaggccaa gacttcaggc atctgggggc agggctattc ccgggccagg 780
ggcacccctgc tgatccactc agtgctgac acattgtacg tgagcacagg ggtggtgttc 840
tccctggaca tgggtgctgac caggtaccac cacattgact ctgggactca cacatggctc 900
ctggcagcta acagtgaggt actcatgatg cttccccgtg ccatgctccc atacctgtac 960
ctgctccgct accggcagct gttgggcatg gtccggggcc acctcccatc caggaggcac 1020
caggccatct ttaccatttc ctag 1044

```

<210> 38  
 <211> 347  
 <212> PRT  
 <213> Homo sapiens

<400> 38

```

Met Gly Asp Glu Leu Ala Pro Cys Pro Val Gly Thr Thr Ala Trp Pro
1          5          10          15
Ala Leu Ile Gln Leu Ile Ser Lys Thr Pro Cys Met Pro Gln Ala Ala
20          25          30
Ser Asn Thr Ser Leu Gly Leu Gly Asp Leu Arg Val Pro Ser Ser Met
35          40          45
Leu Tyr Trp Leu Phe Leu Pro Ser Ser Leu Leu Ala Ala Ala Thr Leu
50          55          60
Ala Val Ser Pro Leu Leu Leu Val Thr Ile Leu Arg Asn Gln Arg Leu
65          70          75          80
Arg Gln Glu Pro His Tyr Leu Leu Pro Ala Asn Ile Leu Leu Ser Asp
85          90          95
Leu Ala Tyr Ile Leu Leu His Met Leu Ile Ser Ser Ser Ser Leu Gly
100         105         110
Gly Trp Glu Leu Gly Arg Met Ala Cys Gly Ile Leu Thr Asp Ala Val
115         120         125
Phe Ala Ala Cys Thr Ser Thr Ile Leu Ser Phe Thr Ala Ile Val Leu
130         135         140
His Thr Tyr Leu Ala Val Ile His Pro Leu Arg Tyr Leu Ser Phe Met
145         150         155         160
Ser His Gly Ala Ala Trp Lys Ala Val Ala Leu Ile Trp Leu Val Ala
165         170         175

```

Cys Cys Phe Pro Thr Phe Leu Ile Trp Leu Ser Lys Trp Gln Asp Ala  
 180 185 190  
 Gln Leu Glu Glu Gln Gly Ala Ser Tyr Ile Leu Pro Pro Ser Met Gly  
 195 200 205  
 Thr Gln Pro Gly Cys Gly Leu Leu Val Ile Val Thr Tyr Thr Ser Ile  
 210 215 220  
 Leu Cys Val Leu Phe Leu Cys Thr Ala Leu Ile Ala Asn Cys Phe Trp  
 225 230 235 240  
 Arg Ile Tyr Ala Glu Ala Lys Thr Ser Gly Ile Trp Gly Gln Gly Tyr  
 245 250 255  
 Ser Arg Ala Arg Gly Thr Leu Leu Ile His Ser Val Leu Ile Thr Leu  
 260 265 270  
 Tyr Val Ser Thr Gly Val Val Phe Ser Leu Asp Met Val Leu Thr Arg  
 275 280 285  
 Tyr His His Ile Asp Ser Gly Thr His Thr Trp Leu Leu Ala Ala Asn  
 290 295 300  
 Ser Glu Val Leu Met Met Leu Pro Arg Ala Met Leu Pro Tyr Leu Tyr  
 305 310 315 320  
 Leu Leu Arg Tyr Arg Gln Leu Leu Gly Met Val Arg Gly His Leu Pro  
 325 330 335  
 Ser Arg Arg His Gln Ala Ile Phe Thr Ile Ser  
 340 345

<210> 39  
 <211> 1023  
 <212> DNA  
 <213> Homo sapiens

<400> 39  
 atgaatccat ttcatgcac ttgttgaac acctctgccg aactttttaa caaatcctgg 60  
 aataaagagt ttgcttatca aactgccagt gtggtagata cagtcaccc ccttccatg 120  
 attgggatta tctgttcaac agggctggtt ggcaacatcc tcattgtatt cactataata 180  
 agatccagga aaaaaacagt ccctgacatc tatatctgca acctggctgt ggctgatttg 240  
 gtccacatag ttggaatgcc ttttcttatt caccaatggg cccgaggggg agagtgggtg 300  
 tttggggggc ctctctgcac catcatcaca tccctggata ctgtgaacca atttgccctg 360  
 agtgccatca tgactgtaat gagtgtggac aggtactttg ccctcgtcca accatttcga 420  
 ctgacacgtt ggagaacaag gtacaagacc atccggatca atttgggcct ttgggcagct 480  
 tcctttatcc tggcattgcc tgtctgggtc tactcgaagg tcatcaaatt taaagacggt 540  
 gttgagagtt gtgcttttga ttgacatcc cctgacgatg tactctggta tacactttat 600  
 ttgacgataa caactttttt tttccctcta cccttgattt tgggtgtgcta tattttaatt 660  
 ttatgctata cttgggagat gtatcaacag aataaggatg ccagatgctg caatcccagt 720  
 gtacaaaaac agagagtgat gaagttgaca aagatggtgc tgggtctggt ggtagtcttt 780

```

atcctgagtg ctgccccctta tcatgtgata caactgggtga acttacagat ggaacagccc 840
acactggcct tctatgtggg ttattacctc tccatctgtc tcagctatgc cagcagcagc 900
attaaccctt ttctctacat cctgctgagt ggaaatttcc agaaacgtct gcctcaaatc 960
caaagaagag cgactgagaa ggaaatcaac aatatgggaa acactctgaa atcacacttt 1020
tag 1023

```

```

<210> 40
<211> 340
<212> PRT
<213> Homo sapiens

```

```

<400> 40

```

```

Met Asn Pro Phe His Ala Ser Cys Trp Asn Thr Ser Ala Glu Leu Leu
1          5          10          15
Asn Lys Ser Trp Asn Lys Glu Phe Ala Tyr Gln Thr Ala Ser Val Val
20          25          30
Asp Thr Val Ile Leu Pro Ser Met Ile Gly Ile Ile Cys Ser Thr Gly
35          40          45
Leu Val Gly Asn Ile Leu Ile Val Phe Thr Ile Ile Arg Ser Arg Lys
50          55          60
Lys Thr Val Pro Asp Ile Tyr Ile Cys Asn Leu Ala Val Ala Asp Leu
65          70          75          80
Val His Ile Val Gly Met Pro Phe Leu Ile His Gln Trp Ala Arg Gly
85          90          95
Gly Glu Trp Val Phe Gly Gly Pro Leu Cys Thr Ile Ile Thr Ser Leu
100         105         110
Asp Thr Cys Asn Gln Phe Ala Cys Ser Ala Ile Met Thr Val Met Ser
115         120         125
Val Asp Arg Tyr Phe Ala Leu Val Gln Pro Phe Arg Leu Thr Arg Trp
130         135         140
Arg Thr Arg Tyr Lys Thr Ile Arg Ile Asn Leu Gly Leu Trp Ala Ala
145         150         155         160
Ser Phe Ile Leu Ala Leu Pro Val Trp Val Tyr Ser Lys Val Ile Lys
165         170         175
Phe Lys Asp Gly Val Glu Ser Cys Ala Phe Asp Leu Thr Ser Pro Asp
180         185         190
Asp Val Leu Trp Tyr Thr Leu Tyr Leu Thr Ile Thr Thr Phe Phe Phe
195         200         205
Pro Leu Pro Leu Ile Leu Val Cys Tyr Ile Leu Ile Leu Cys Tyr Thr
210         215         220
Trp Glu Met Tyr Gln Gln Asn Lys Asp Ala Arg Cys Cys Asn Pro Ser
225         230         235         240
Val Pro Lys Gln Arg Val Met Lys Leu Thr Lys Met Val Leu Val Leu

```

245                      250                      255  
 Val Val Val Phe Ile Leu Ser Ala Ala Pro Tyr His Val Ile Gln Leu  
                  260                      265                      270  
 Val Asn Leu Gln Met Glu Gln Pro Thr Leu Ala Phe Tyr Val Gly Tyr  
                  275                      280                      285  
 Tyr Leu Ser Ile Cys Leu Ser Tyr Ala Ser Ser Ser Ile Asn Pro Phe  
                  290                      295                      300  
 Leu Tyr Ile Leu Leu Ser Gly Asn Phe Gln Lys Arg Leu Pro Gln Ile  
                  305                      310                      315                      320  
 Gln Arg Arg Ala Thr Glu Lys Glu Ile Asn Asn Met Gly Asn Thr Leu  
                  325                      330                      335  
 Lys Ser His Phe  
                  340

<210> 41  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 41  
 cttgcagaca tcaccatggc agcc

24

<210> 42  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 42  
 gtgatgctct gagtactgga ctgg

24

<210> 43  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 43  
 gaagctgtga agagtgtgc

20

<210> 44  
 <211> 24  
 <212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 44

gtcagcaata ttgataagca gcag

24

<210> 45

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 45

ccatggggaa cgattctgtc agctacg

27

<210> 46

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 46

gctatgcctg aagccagtct tgtg

24

<210> 47

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 47

ccaggatggt gtgtcaccgt ggtggc

26

<210> 48

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 48

cacagcgtg cagccctgca gctggc

26

<210> 49  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 49  
cttcctctcg tagggatgaa ccagac

26

<210> 50  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 50  
ctcgcacagg tgggaagcac ctgtgg

26

<210> 51  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 51  
gcctgtgaca ggaggtaccc tgg

23

<210> 52  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 52  
catatccctc cgagtgtcca gcggc

25

<210> 53  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature

<223> Novel Sequence

<400> 53  
gcatggagag aaaatttatg tccttgcaac c

31

<210> 54  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 54  
caagaacagg tctcatctaa gagctcc

27

<210> 55  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 55  
gctgttgcca tgacgtccac ctgcac

26

<210> 56  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 56  
ggacagttca aggtttgcct tagaac

26

<210> 57  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 57  
ctttcgatac tgctcctatg ctc

23

<210> 58  
<211> 26



<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 58  
gtagtccact gaaagtccag tgatcc

26

<210> 59  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 59  
tttctgagca tggatccaac catctc

26

<210> 60  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 60  
ctgtctgaca gggcagaggc tcttc

25

<210> 61  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 61  
ggaactcgta tagaccagc gtcgctcc

28

<210> 62  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 62

ggaggttgcg ccttagcgac agatgacc

28

<210> 63  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 63  
ctgcaccg acacttgctc tg

22

<210> 64  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 64  
gtctgcttgt tcagtggcac tcaac

25

<210> 65  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 65  
tatctgcaat tctattctag ctctg

26

<210> 66  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 66  
tgtccctaataaagtcacat gaatgc

26

<210> 67  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>

<221> misc\_feature  
<223> Novel Sequence

<400> 67  
ggagacaacc atgaatgagc cac

23

<210> 68  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 68  
tatttcaagg gttgtttgag taac

24

<210> 69  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 69  
ggcaccagtg gaggttttct gagcatg

27

<210> 70  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 70  
ctgatggaag tagaggctgt ccatctc

27

<210> 71  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 71  
cctggcgagc cgctagcgcc atg

23

<210> 72

<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 72  
atgagccctg ccaggccctc agt

23

<210> 73  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 73  
ctgcgatgcc cacactcaat acttctg

27

<210> 74  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 74  
aaggatccta cacttggtgg atctcag

27

<210> 75  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<400> 75  
gctggagcat tcactaggcg ag

22

<210> 76  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 76  
agatcctggt tcttggtgac aatg

24

<210> 77

<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 77  
agccatccct gccaggaagc atgg

24

<210> 78  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 78  
ccagactgtg gactcaagaa ctctagg

27

<210> 79  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 79  
agtccacgaa caatgaatcc atttcatg

28

<210> 80  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 80  
atcatgtcta gactcatggt gatcc

25

<210> 81  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 81  
ggggagggaag agcaaagggtg gtcctcctgg 30

<210> 82  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 82  
ccaggagaac cacctttgct ttcctccccc 30

<210> 83  
<211> 1356  
<212> DNA  
<213> Homo sapiens

<400> 83  
atggagtcct caccatccc ccagtcacga gggaactctt ccactttggg gaggggtccct 60  
caaacccccag gtccctctac tgccagtggt gtcccgagg tggggctacg ggatgttgct 120  
tcggaatctg tggccctctt cttcatgctc ctgctggact tgactgctgt ggctggcaat 180  
gccgctgtga tggccgtgat cgccaagacg cctgccctcc gaaaatttgt cttcgtcttc 240  
cacctctgcc tgggtggacct gctggctgcc ctgacctca tgcccctggc catgctctcc 300  
agctctgccc tctttgacca cgccctcttt ggggaggtgg cctgccgcct ctacttgttt 360  
ctgagcgtgt gctttgtcag cctggccatc ctctcgtgt cagccatcaa tgtggagcgc 420  
tactattacg tagtccaccc catgcgtac gaggtgcga tgacgtggg gctgggtggc 480  
tctgtgctgg tgggtgtgtg ggtgaaggcc ttggccatgg cttctgtgcc agtgttggga 540  
aggtctcctt gggaggaagg agtcccagt gtccccccag gctgttcaact ccagtggagc 600  
cacagtgcct actgccagct ttttgtggtg gtctttgctg tcctttactt tctgttgccc 660  
ctgtcctca tacttgtggt ctactgcagc atgttccgag tggccgcgt ggctgccatg 720  
cagcacgggc cgctgccac gtgatggag acacccggc aacgctccga atctctcagc 780  
agccgctcca cgatgggtcac cagctcgggg gccccccaga ccacccaca ccggacgttt 840  
gggggaggga aagcaaagggt ggttctcctg gctgtggggg gacagtccct gctctgttgg 900  
ttgccctact tctctttcca cctctatgtt gccctgagtg ctacgcccct ttcaactggg 960  
caggtggaga gtgtggtcac ctggattggc tacttttgct tcacttccaa ccctttcttc 1020  
tatggatgtc tcaaccggca gatccggggg gagctcagca agcagtttgt ctgcttcttc 1080  
aagccagctc cagaggagga gctgaggctg cctagccggg agggctccat tgaggagaac 1140  
ttcctgcagt tccttcaggg gactggctgt ccttctgagt cctgggtttc ccgacccta 1200  
cccagccca agcaggagcc acctgctgtt gactttcgaa tcccaggcca gatagctgag 1260

gagacctctg agttcctgga gcagcaactc accagcgaca tcatcatgtc agacagctac 1320  
 ctccgtcctg ccgcctcacc ccggctggag tcatga 1356

<210> 84  
 <211> 451  
 <212> PRT  
 <213> Homo sapiens

<400> 84

Met Glu Ser Ser Pro Ile Pro Gln Ser Ser Gly Asn Ser Ser Thr Leu  
 1 5 10 15  
 Gly Arg Val Pro Gln Thr Pro Gly Pro Ser Thr Ala Ser Gly Val Pro  
 20 25 30  
 Glu Val Gly Leu Arg Asp Val Ala Ser Glu Ser Val Ala Leu Phe Phe  
 35 40 45  
 Met Leu Leu Leu Asp Leu Thr Ala Val Ala Gly Asn Ala Ala Val Met  
 50 55 60  
 Ala Val Ile Ala Lys Thr Pro Ala Leu Arg Lys Phe Val Phe Val Phe  
 65 70 75 80  
 His Leu Cys Leu Val Asp Leu Leu Ala Ala Leu Thr Leu Met Pro Leu  
 85 90 95  
 Ala Met Leu Ser Ser Ser Ala Leu Phe Asp His Ala Leu Phe Gly Glu  
 100 105 110  
 Val Ala Cys Arg Leu Tyr Leu Phe Leu Ser Val Cys Phe Val Ser Leu  
 115 120 125  
 Ala Ile Leu Ser Val Ser Ala Ile Asn Val Glu Arg Tyr Tyr Tyr Val  
 130 135 140  
 Val His Pro Met Arg Tyr Glu Val Arg Met Thr Leu Gly Leu Val Ala  
 145 150 155 160  
 Ser Val Leu Val Gly Val Trp Val Lys Ala Leu Ala Met Ala Ser Val  
 165 170 175  
 Pro Val Leu Gly Arg Val Ser Trp Glu Glu Gly Ala Pro Ser Val Pro  
 180 185 190  
 Pro Gly Cys Ser Leu Gln Trp Ser His Ser Ala Tyr Cys Gln Leu Phe  
 195 200 205  
 Val Val Val Phe Ala Val Leu Tyr Phe Leu Leu Pro Leu Leu Leu Ile  
 210 215 220  
 Leu Val Val Tyr Cys Ser Met Phe Arg Val Ala Arg Val Ala Ala Met  
 225 230 235 240  
 Gln His Gly Pro Leu Pro Thr Trp Met Glu Thr Pro Arg Gln Arg Ser  
 245 250 255  
 Glu Ser Leu Ser Ser Arg Ser Thr Met Val Thr Ser Ser Gly Ala Pro  
 260 265 270  
 Gln Thr Thr Pro His Arg Thr Phe Gly Gly Gly Lys Ala Lys Val Val

275	280	285	
Leu Leu Ala Val Gly Gly Gln Phe Leu Leu Cys Trp Leu Pro Tyr Phe 290 295 300			
Ser Phe His Leu Tyr Val Ala Leu Ser Ala Gln Pro Ile Ser Thr Gly 305 310 315 320			
Gln Val Glu Ser Val Val Thr Trp Ile Gly Tyr Phe Cys Phe Thr Ser 325 330 335			
Asn Pro Phe Phe Tyr Gly Cys Leu Asn Arg Gln Ile Arg Gly Glu Leu 340 345 350			
Ser Lys Gln Phe Val Cys Phe Phe Lys Pro Ala Pro Glu Glu Glu Leu 355 360 365			
Arg Leu Pro Ser Arg Glu Gly Ser Ile Glu Glu Asn Phe Leu Gln Phe 370 375 380			
Leu Gln Gly Thr Gly Cys Pro Ser Glu Ser Trp Val Ser Arg Pro Leu 385 390 395 400			
Pro Ser Pro Lys Gln Glu Pro Pro Ala Val Asp Phe Arg Ile Pro Gly 405 410 415			
Gln Ile Ala Glu Glu Thr Ser Glu Phe Leu Glu Gln Gln Leu Thr Ser 420 425 430			
Asp Ile Ile Met Ser Asp Ser Tyr Leu Arg Pro Ala Ala Ser Pro Arg 435 440 445			
Leu Glu Ser 450			
<210> 85			
<211> 28			
<212> DNA			
<213> Homo sapiens			
<400> 85			
caggaaggca aagaccacca tcatcatc			28
<210> 86			
<211> 28			
<212> DNA			
<213> Homo sapiens			
<400> 86			
gatgatgatg gtggtctttg ccttcctg			28
<210> 87			
<211> 1041			
<212> DNA			
<213> Homo sapiens			
<400> 87			
atggagagaa aatttatgtc cttgcaacca tccatctccg tatcagaaat ggaaccaa			60
ggcaccttca gcaataacaa cagcaggaac tgcacaattg aaaacttcaa gagagaattt			120
ttccaattg tatatctgat aatatttttc tggggagtct tgggaaatgg gttgtccata			180



```

tatgttttcc tgcagcctta taagaagtcc acatctgtga acgttttcat gctaaatctg 240
gccatttcag atctcctgtt cataagcacg cttcccttca gggctgacta ttatcttaga 300
ggctccaatt ggatatttgg agacctggcc tgcaggatta tgtcttattc cttgtatgtc 360
aacatgtaca gcagtattta tttcctgacc gtgctgagtg ttgtgcgttt cctggcaatg 420
gttcaccctt ttcggcttct gcatgtcacc agcatcagga gtgcctggat cctctgtggg 480
atcatatgga tccttatcat ggcttcctca ataatgctcc tggacagtgg ctctgagcag 540
aacggcagtg tcacatcatg cttagagctg aatctctata aaattgctaa gctgcagacc 600
atgaactata ttgccttggg ggtgggctgc ctgctgccat tttcacact cagcatctgt 660
tatctgctga tcattcgggt tctgttaaaa gtggaggtcc cagaatcggg gctgcggggt 720
tctcacagga aggcaaagac caccatcatc atcaccttga tcatcttctt cttgtgtttc 780
ctgccttatc acacactgag gaccgtccac ttgacgacat ggaaagtggg tttatgcaaa 840
gacagactgc ataaagcttt ggttatcaca ctggccttgg cagcagccaa tgcctgcttc 900
aatcctctgc tctattactt tgctggggag aattttaagg acagactaaa gtctgcactc 960
agaaaaggcc atccacagaa ggcaaagaca aagtgtgttt tccctgttag tgtgtggtg 1020
agaaaggaaa caagagtata a 1041

```

<210> 88  
 <211> 346  
 <212> PRT  
 <213> Homo sapiens

<400> 88

```

Met Glu Arg Lys Phe Met Ser Leu Gln Pro Ser Ile Ser Val Ser Glu
1          5          10          15
Met Glu Pro Asn Gly Thr Phe Ser Asn Asn Asn Ser Arg Asn Cys Thr
20          25          30
Ile Glu Asn Phe Lys Arg Glu Phe Phe Pro Ile Val Tyr Leu Ile Ile
35          40          45
Phe Phe Trp Gly Val Leu Gly Asn Gly Leu Ser Ile Tyr Val Phe Leu
50          55          60
Gln Pro Tyr Lys Lys Ser Thr Ser Val Asn Val Phe Met Leu Asn Leu
65          70          75          80
Ala Ile Ser Asp Leu Leu Phe Ile Ser Thr Leu Pro Phe Arg Ala Asp
85          90          95
Tyr Tyr Leu Arg Gly Ser Asn Trp Ile Phe Gly Asp Leu Ala Cys Arg
100         105         110
Ile Met Ser Tyr Ser Leu Tyr Val Asn Met Tyr Ser Ser Ile Tyr Phe
115         120         125
Leu Thr Val Leu Ser Val Val Arg Phe Leu Ala Met Val His Pro Phe
130         135         140

```

Arg Leu Leu His Val Thr Ser Ile Arg Ser Ala Trp Ile Leu Cys Gly  
 145 150 155 160  
 Ile Ile Trp Ile Leu Ile Met Ala Ser Ser Ile Met Leu Leu Asp Ser  
 165 170 175  
 Gly Ser Glu Gln Asn Gly Ser Val Thr Ser Cys Leu Glu Leu Asn Leu  
 180 185 190  
 Tyr Lys Ile Ala Lys Leu Gln Thr Met Asn Tyr Ile Ala Leu Val Val  
 195 200 205  
 Gly Cys Leu Leu Pro Phe Phe Thr Leu Ser Ile Cys Tyr Leu Leu Ile  
 210 215 220  
 Ile Arg Val Leu Leu Lys Val Glu Val Pro Glu Ser Gly Leu Arg Val  
 225 230 235 240  
 Ser His Arg Lys Ala Lys Thr Thr Ile Ile Ile Thr Leu Ile Ile Phe  
 245 250 255  
 Phe Leu Cys Phe Leu Pro Tyr His Thr Leu Arg Thr Val His Leu Thr  
 260 265 270  
 Thr Trp Lys Val Gly Leu Cys Lys Asp Arg Leu His Lys Ala Leu Val  
 275 280 285  
 Ile Thr Leu Ala Leu Ala Ala Ala Asn Ala Cys Phe Asn Pro Leu Leu  
 290 295 300  
 Tyr Tyr Phe Ala Gly Glu Asn Phe Lys Asp Arg Leu Lys Ser Ala Leu  
 305 310 315 320  
 Arg Lys Gly His Pro Gln Lys Ala Lys Thr Lys Cys Val Phe Pro Val  
 325 330 335  
 Ser Val Trp Leu Arg Lys Glu Thr Arg Val  
 340 345

<210> 89  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 89  
 ccagtgc aaa gctaagaaag tgatcttc

28

<210> 90  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 90  
 gaagatcact ttcttagctt tgcactgg

28

<210> 91  
 <211> 1527  
 <212> DNA  
 <213> Homo sapiens

<400> 91  
 atgacgtcca cctgcaccaa cagcacgcgc gagagtaaca gcagccacac gtgcatgccc 60  
 ctctccaaaa tgcccatcag cctggcccac ggcatcatcc gctcaaccgt gctgggtatc 120  
 ttcctcgccg cctctttcgt cggcaacata gtgctggcgc tagtggtgca gcgcaagccg 180  
 cagctgctgc aggtgaccaa ccgttttatc tttaacctcc tcgtcaccga cctgctgcag 240  
 atttcgctcg tggccccctg ggtggtggcc acctctgtgc ctctcttctg gcccccaac 300  
 agccacttct gcacggccct ggtagcctc acccacctgt tcgccttcgc cagcgtcaac 360  
 accattgtcg tgggtgcagt ggatcgctac ttgtccatca tccacctct ctectaccg 420  
 tccaagatga ccagcgccg cggttacctg ctctctatg gcacctggat tgtggccatc 480  
 ctgcagagca ctctccact ctacggctgg ggccaggctg cctttgatga gcgcaatgct 540  
 ctctgctcca tgatctgggg ggccagcccc agctacacta ttctcagcgt ggtgtccttc 600  
 atcgtcattc cactgattgt catgattgcc tgctactccg tgggtgttctg tgcagcccg 660  
 aggcagcatg ctctgctgta caatgtcaag agacacagct tggaagtgcg agtcaaggac 720  
 tgtgtggaga atgaggatga agaggagca gagaagaagg aggagtcca ggatgagagt 780  
 gaggtttcgcc gccagcatga aggtgaggtc aaggccaagg agggcagaat ggaagccaag 840  
 gacggcagcc tgaaggccaa ggaaggaagc acggggacca gtgagagtag tgtagaggcc 900  
 aggggcagcg aggaggtcag agagagcagc acgggtggcca gcgacggcag catggagggt 960  
 aaggaaggca gcaccaaagt tgaaggagaac agcatgaagg cagacaaggg tcgcacagag 1020  
 gtcaaccagt gcagcattga cttgggtgaa gatgacatgg agtttggtga agacgacatc 1080  
 aatttcagtg aggatgacgt cgaggcagt aacatcccgg agagcctccc acccagtcgt 1140  
 cgtaacagca acagcaaccc tcctctgccc aggtgctacc agtgcaaagc taagaaagtg 1200  
 atcttcatca tcattttctc ctatgtgcta tccctggggc cctactgctt tttagcagtc 1260  
 ctggccgtgt ggggtgatgt cgaaacccag gtaccccatg gggatgatcac cataatcatc 1320  
 tggtttttct tcctgcagtg ctgcatccac ccctatgtct atggctacat gcacaagacc 1380  
 attaagaagg aaatccagga catgctgaag aagttcttct gcaaggaaaa gccccgaaa 1440  
 gaagatagcc acccagacct gcccgaaca gaggggggga ctgaaggcaa gattgtccct 1500  
 tcctacgatt ctgctacttt tccttga 1527

<210> 92  
 <211> 508  
 <212> PRT  
 <213> Homo sapiens

&lt;400&gt; 92

```

Met Thr Ser Thr Cys Thr Asn Ser Thr Arg Glu Ser Asn Ser Ser His
1      5      10      15

Thr Cys Met Pro Leu Ser Lys Met Pro Ile Ser Leu Ala His Gly Ile
      20      25      30

Ile Arg Ser Thr Val Leu Val Ile Phe Leu Ala Ala Ser Phe Val Gly
      35      40      45

Asn Ile Val Leu Ala Leu Val Leu Gln Arg Lys Pro Gln Leu Leu Gln
      50      55      60

Val Thr Asn Arg Phe Ile Phe Asn Leu Leu Val Thr Asp Leu Leu Gln
      65      70      75      80

Ile Ser Leu Val Ala Pro Trp Val Val Ala Thr Ser Val Pro Leu Phe
      85      90      95

Trp Pro Leu Asn Ser His Phe Cys Thr Ala Leu Val Ser Leu Thr His
      100      105      110

Leu Phe Ala Phe Ala Ser Val Asn Thr Ile Val Val Val Ser Val Asp
      115      120      125

Arg Tyr Leu Ser Ile Ile His Pro Leu Ser Tyr Pro Ser Lys Met Thr
      130      135      140

Gln Arg Arg Gly Tyr Leu Leu Leu Tyr Gly Thr Trp Ile Val Ala Ile
      145      150      155      160

Leu Gln Ser Thr Pro Pro Leu Tyr Gly Trp Gly Gln Ala Ala Phe Asp
      165      170      175

Glu Arg Asn Ala Leu Cys Ser Met Ile Trp Gly Ala Ser Pro Ser Tyr
      180      185      190

Thr Ile Leu Ser Val Val Ser Phe Ile Val Ile Pro Leu Ile Val Met
      195      200      205

Ile Ala Cys Tyr Ser Val Val Phe Cys Ala Ala Arg Arg Gln His Ala
      210      215      220

Leu Leu Tyr Asn Val Lys Arg His Ser Leu Glu Val Arg Val Lys Asp
      225      230      235      240

Cys Val Glu Asn Glu Asp Glu Glu Gly Ala Glu Lys Lys Glu Glu Phe
      245      250      255

Gln Asp Glu Ser Glu Phe Arg Arg Gln His Glu Gly Glu Val Lys Ala
      260      265      270

Lys Glu Gly Arg Met Glu Ala Lys Asp Gly Ser Leu Lys Ala Lys Glu
      275      280      285

Gly Ser Thr Gly Thr Ser Glu Ser Ser Val Glu Ala Arg Gly Ser Glu
      290      295      300

Glu Val Arg Glu Ser Ser Thr Val Ala Ser Asp Gly Ser Met Glu Gly
      305      310      315      320

Lys Glu Gly Ser Thr Lys Val Glu Glu Asn Ser Met Lys Ala Asp Lys
      325      330      335

```

Gly Arg Thr Glu Val Asn Gln Cys Ser Ile Asp Leu Gly Glu Asp Asp  
 340 345 350  
 Met Glu Phe Gly Glu Asp Asp Ile Asn Phe Ser Glu Asp Asp Val Glu  
 355 360 365  
 Ala Val Asn Ile Pro Glu Ser Leu Pro Pro Ser Arg Arg Asn Ser Asn  
 370 375 380  
 Ser Asn Pro Pro Leu Pro Arg Cys Tyr Gln Cys Lys Ala Lys Lys Val  
 385 390 395 400  
 Ile Phe Ile Ile Ile Phe Ser Tyr Val Leu Ser Leu Gly Pro Tyr Cys  
 405 410 415  
 Phe Leu Ala Val Leu Ala Val Trp Val Asp Val Glu Thr Gln Val Pro  
 420 425 430  
 Gln Trp Val Ile Thr Ile Ile Ile Trp Leu Phe Phe Leu Gln Cys Cys  
 435 440 445  
 Ile His Pro Tyr Val Tyr Gly Tyr Met His Lys Thr Ile Lys Lys Glu  
 450 455 460  
 Ile Gln Asp Met Leu Lys Lys Phe Phe Cys Lys Glu Lys Pro Pro Lys  
 465 470 475 480  
 Glu Asp Ser His Pro Asp Leu Pro Gly Thr Glu Gly Gly Thr Glu Gly  
 485 490 495  
 Lys Ile Val Pro Ser Tyr Asp Ser Ala Thr Phe Pro  
 500 505

<210> 93  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 93.  
 gccgccaccg cgccaagagg aagattggc

29

<210> 94  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 94  
 gccaatcttc ctcttggcgc ggtggcggc

29

<210> 95  
 <211> 1092  
 <212> DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 95

```

atgggccccg gcgaggcgct gctggcgggg ctccctggtga tgggtactggc cgtggcgctg      60
ctatccaacg cactggtgct gctttgttgc gcctacagcg ctgagctccg cactcgagcc      120
tcaggcgctc tcctggtgaa tctgtcgctg gggcacctgc tgctggcggc gctggacatg      180
cccttcacgc tgctcggtgt gatgcgcggg cggacaccgt cggcgcccgg cgcattgcaa      240
gtcattggct tcctggacac cttcctggcg tccaacgcgg cgctgagcgt ggcggcgctg      300
agcgagagcc agtggctggc agtgggcttc ccactgcgct acgccggacg cctgcgaccg      360
cgctatgccg gcctgctgct gggctgtgcc tggggacagt cgctggcctt ctcaggcgct      420
gcacttggct gctcgtggct tggctacagc agcgcccttc cgctcctgttc gctgcgcctg      480
ccgcccagag ctgagcgctc gcgcttcgca gccttcaccg ccacgctcca tgccgtgggc      540
ttcgtgctgc cgctggcggt gctctgcctc acctcgctcc aggtgcaccg ggtggcacgc      600
agccactgcc agcgcatgga caccgtcacc atgaaggcgc tcgcgctgct cgccgacctg      660
caccacagtg tgcggcagcg ctgcctcatc cagcagaagc ggcgccgcca ccgcgccacc      720
aggaagattg gcattgctat tgcgaccttc ctcatctgct ttgccccgta tgtcatgacc      780
aggctggcgg agctcgtgcc cttcgtcacc gtgaacgccc agaaggcat cctcagcaag      840
tgcctgacct acagcaaggc ggtggccgac ccgttcacgt actctctgct ccgcccggccg      900
ttccgccaa gtcctggccgg catggtgcac cggtgctga agagaacccc gcgcccagca      960
tccacccatg acagctctct ggatgtggcc ggcattggtgc accagctgct gaagagaacc     1020
ccgcgcccag cgtccaccga caacggctct gtggacacag agaatgattc ctgcctgcag     1080
cagacacact ga                                           1092

```

&lt;210&gt; 96

&lt;211&gt; 363

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 96

```

Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met Val Leu
1          5          10          15
Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr
20          25          30
Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu
35          40          45
Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu
50          55          60
Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln
65          70          75          80
Val Ile Gly Phe Leu Asp Thr Ph Leu Ala Ser Asn Ala Ala Leu Ser

```

<400> 97  
gatctctaga atggagtcct cacccatccc ccag

<210> 98  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 98  
 gatcgatatc cgtgactcca gccgggggtga ggcggc 36

<210> 99  
 <211> 2610  
 <212> DNA  
 <213> Homo sapiens and Rat

<400> 99  
 atggagtcct caccatccc ccagtcacga gggaaactctt ccactttggg gagggtcctt 60  
 caaaccccag gtccctctac tgccagtggg gtcccgaggg tggggctacg ggatgttgct 120  
 tcggaatctg tggccctctt ctcatgctc ctgctggact tgactgctgt ggctggcaat 180  
 gccgctgtga tggccgtgat cgccaagacg cctgccctcc gaaaatttgt cttcgtcttc 240  
 cacctctgcc tgggtggacct gctggctgcc ctgaccctca tgccctggc catgctctcc 300  
 agctctgccc tctttgacca cgccctcttt ggggaggtgg cctgccgcct ctacttgttt 360  
 ctgagcgtgt gctttgtcag cctggccatc ctctcgggtg cagccatcaa tgtggagcgc 420  
 tactattacg tagtccaccc catgcgctac gaggtgcgca tgacgctggg gctggtggcc 480  
 tctgtgctgg tgggtgtgtg ggtgaaggcc ttggccatgg cttctgtgcc agtgttggga 540  
 agggctctct gggaggaagg agctcccagt gtccccccag gctgttact ccagtggagc 600  
 cacagtgcct actgccagct ttttgtggtg gtctttgtct tcctttactt tctgttgccc 660  
 ctgctcctca tacttgtggt ctactgcagc atgttccgag tggccgcgt ggctgccatg 720  
 cagcacgggc cgctgcccac gtggatggag acacccggc aacgctccga atctctcagc 780  
 agccgctcca cgatggtcac cagctcgggg gccccccaga ccacccaca ccggacgttt 840  
 gggggaggga aagcagcagt ggttctctg gctgtggggg gacagttcct gctctgttgg 900  
 ttgccctact tctctttcca cctctatgtt gccctgagt ctcagcccat ttcaactggg 960  
 cagggtggaga gtgtggtcac ctggattggc tacttttgct tcacttccaa ccctttcttc 1020  
 tatggatgtc tcaaccggca gatccggggg gagctcagca agcagtttgt ctgcttcttc 1080  
 aagccagctc cagaggagga gctgaggctg cctagccggg agggctccat tgaggagaac 1140  
 ttctctcagt tccttcaggg gactggctgt ccttctgagt cctgggtttc ccgacccta 1200  
 cccagcccca agcaggagcc acctgctgtt gactttcgaa tcccaggcca gatagctgag 1260  
 gagacctctg agttcctgga gcagcaactc accagcgaca tcatcatgtc agacagctac 1320



```

ctccgtcctg cgcctcacc ccggttgag tcagcgatat ctgcagaatt ccaccacact 1380
ggactagtgg atccgagctc ggtaccaagc ttgggctgca ggtcgatggg ctgcctcggc 1440
aacagtaaga ccgaggacca gcgcaacgag gagaaggcgc agcgcgaggc caacaaaaag 1500
atcgagaagc agctgcagaa ggacaagcag gtctaccggg ccacgcaccg cctgctgctg 1560
ctgggtgctg gagagtctgg caaaagcacc attgtgaagc agatgaggat cctacatggt 1620
aatgggttta acggagaggg cggcgaagag gacccgcagg ctgcaaggag caacagcgat 1680
ggtgagaagg ccaccaaagt gcaggacatc aaaaacaacc tgaaggaggc cattgaaacc 1740
attgtggccg ccatgagcaa cctggtgccc cccgtggagc tggccaaccc tgagaaccag 1800
ttcagagtgg actacattct gagcgtgatg aacgtgcaa actttgactt cccacctgaa 1860
ttctatgagc atgccaaggc tctgtgggag gatgaggagg ttcgtgcctg ctacgagcgc 1920
tccaacgagt accagctgat cgactgtgcc cagtacttcc tggacaagat tgatgtgatc 1980
aagcaggccg actacgtgcc aagtgaccag gacctgcttc gctgccgcgt cctgacctct 2040
ggaatctttg agaccaagtt ccaggtggac aaagtcaact tccacatggt cgatgtgggc 2100
ggccagcgcg atgaacgccg caagtggatc cagtgttca atgatgtgac tgccatcatc 2160
ttcgtggtgg ccagcagcag ctacaacatg gtcacccggg aggacaacca gaccaaccgt 2220
ctgcaggagg ctctgaacct cttcaagagc atctggaaca acagatggct gcgtaccatc 2280
tctgtgatcc tcttctcaa caagcaagat ctgcttgctg agaaggctct cgctgggaaa 2340
tcgaagattg aggactactt tccagagttc gctcgtaca ccactcctga ggatgcgact 2400
cccagccccg gagaggaccc acgcgtgacc cgggccaaagt acttcatccg ggatgagttt 2460
ctgagaatca gcactgctag tggagatgga cgctactact gctaccctca ctttacctgc 2520
gccgtggaca ctgagaacat ccgccgtgtc ttcaacgact gccgtgacat catccagcgc 2580
atgcatcttc gccaatacga gctgctctaa 2610

```

&lt;210&gt; 100

&lt;211&gt; 869

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens and Rat

&lt;400&gt; 100

```

Met Glu Ser Ser Pro Ile Pro Gln Ser Ser Gly Asn Ser Ser Thr Leu
1           5           10           15

```

```

Gly Arg Val Pro Gln Thr Pro Gly Pro Ser Thr Ala Ser Gly Val Pro
20           25           30

```

```

Glu Val Gly Leu Arg Asp Val Ala Ser Glu Ser Val Ala Leu Phe Phe
35           40           45

```

```

Met Leu Leu Leu Asp Leu Thr Ala Val Ala Gly Asn Ala Ala Val Met
50           55           60

```

Ala Val Ile Ala Lys Thr Pro Ala Leu Arg Lys Phe Val Phe Val Phe  
 65 70 75 80  
 His Leu Cys Leu Val Asp Leu Leu Ala Ala Leu Thr Leu Met Pro Leu  
 85 90 95  
 Ala Met Leu Ser Ser Ser Ala Leu Phe Asp His Ala Leu Phe Gly Glu  
 100 105 110  
 Val Ala Cys Arg Leu Tyr Leu Phe Leu Ser Val Cys Phe Val Ser Leu  
 115 120 125  
 Ala Ile Leu Ser Val Ser Ala Ile Asn Val Glu Arg Tyr Tyr Tyr Val  
 130 135 140  
 Val His Pro Met Arg Tyr Glu Val Arg Met Thr Leu Gly Leu Val Ala  
 145 150 155 160  
 Ser Val Leu Val Gly Val Trp Val Lys Ala Leu Ala Met Ala Ser Val  
 165 170 175  
 Pro Val Leu Gly Arg Val Ser Trp Glu Glu Gly Ala Pro Ser Val Pro  
 180 185 190  
 Pro Gly Cys Ser Leu Gln Trp Ser His Ser Ala Tyr Cys Gln Leu Phe  
 195 200 205  
 Val Val Val Phe Ala Val Leu Tyr Phe Leu Leu Pro Leu Leu Leu Ile  
 210 215 220  
 Leu Val Val Tyr Cys Ser Met Phe Arg Val Ala Arg Val Ala Ala Met  
 225 230 235 240  
 Gln His Gly Pro Leu Pro Thr Trp Met Glu Thr Pro Arg Gln Arg Ser  
 245 250 255  
 Glu Ser Leu Ser Ser Arg Ser Thr Met Val Thr Ser Ser Gly Ala Pro  
 260 265 270  
 Gln Thr Thr Pro His Arg Thr Phe Gly Gly Gly Lys Ala Ala Val Val  
 275 280 285  
 Leu Leu Ala Val Gly Gly Gln Phe Leu Leu Cys Trp Leu Pro Tyr Phe  
 290 295 300  
 Ser Phe His Leu Tyr Val Ala Leu Ser Ala Gln Pro Ile Ser Thr Gly  
 305 310 315 320  
 Gln Val Glu Ser Val Val Thr Trp Ile Gly Tyr Phe Cys Phe Thr Ser  
 325 330 335  
 Asn Pro Phe Phe Tyr Gly Cys Leu Asn Arg Gln Ile Arg Gly Glu Leu  
 340 345 350  
 Ser Lys Gln Phe Val Cys Phe Phe Lys Pro Ala Pro Glu Glu Glu Leu  
 355 360 365  
 Arg Leu Pro Ser Arg Glu Gly Ser Ile Glu Glu Asn Phe Leu Gln Phe  
 370 375 380  
 Leu Gln Gly Thr Gly Cys Pro Ser Glu Ser Trp Val Ser Arg Pro Leu  
 385 390 395 400  
 Pro Ser Pro Lys Gln Glu Pro Pro Ala Val Asp Phe Arg Ile Pro Gly  
 405 410 415

Gln Ile Ala Glu Glu Thr Ser Glu Phe Leu Glu Gln Gln Leu Thr Ser  
 420 425 430  
 Asp Ile Ile Met Ser Asp Ser Tyr Leu Arg Pro Ala Ala Ser Pro Arg  
 435 440 445  
 Leu Glu Ser Ala Ile Ser Ala Glu Phe His His Thr Gly Leu Val Asp  
 450 455 460  
 Pro Ser Ser Val Pro Ser Leu Gly Cys Arg Ser Met Gly Cys Leu Gly  
 465 470 475 480  
 Asn Ser Lys Thr Glu Asp Gln Arg Asn Glu Glu Lys Ala Gln Arg Glu  
 485 490 495  
 Ala Asn Lys Lys Ile Glu Lys Gln Leu Gln Lys Asp Lys Gln Val Tyr  
 500 505 510  
 Arg Ala Thr His Arg Leu Leu Leu Leu Gly Ala Gly Glu Ser Gly Lys  
 515 520 525  
 Ser Thr Ile Val Lys Gln Met Arg Ile Leu His Val Asn Gly Phe Asn  
 530 535 540  
 Gly Glu Gly Gly Glu Glu Asp Pro Gln Ala Ala Arg Ser Asn Ser Asp  
 545 550 555 560  
 Gly Glu Lys Ala Thr Lys Val Gln Asp Ile Lys Asn Asn Leu Lys Glu  
 565 570 575  
 Ala Ile Glu Thr Ile Val Ala Ala Met Ser Asn Leu Val Pro Pro Val  
 580 585 590  
 Glu Leu Ala Asn Pro Glu Asn Gln Phe Arg Val Asp Tyr Ile Leu Ser  
 595 600 605  
 Val Met Asn Val Pro Asn Phe Asp Phe Pro Pro Glu Phe Tyr Glu His  
 610 615 620  
 Ala Lys Ala Leu Trp Glu Asp Glu Gly Val Arg Ala Cys Tyr Glu Arg  
 625 630 635 640  
 Ser Asn Glu Tyr Gln Leu Ile Asp Cys Ala Gln Tyr Phe Leu Asp Lys  
 645 650 655  
 Ile Asp Val Ile Lys Gln Ala Asp Tyr Val Pro Ser Asp Gln Asp Leu  
 660 665 670  
 Leu Arg Cys Arg Val Leu Thr Ser Gly Ile Phe Glu Thr Lys Phe Gln  
 675 680 685  
 Val Asp Lys Val Asn Phe His Met Phe Asp Val Gly Gly Gln Arg Asp  
 690 695 700  
 Glu Arg Arg Lys Trp Ile Gln Cys Phe Asn Asp Val Thr Ala Ile Ile  
 705 710 715 720  
 Phe Val Val Ala Ser Ser Ser Tyr Asn Met Val Ile Arg Glu Asp Asn  
 725 730 735  
 Gln Thr Asn Arg Leu Gln Glu Ala Leu Asn Leu Phe Lys Ser Ile Trp  
 740 745 750  
 Asn Asn Arg Trp Leu Arg Thr Ile Ser Val Ile Leu Phe Leu Asn Lys

755                      760                      765  
 Gln Asp Leu Leu Ala Glu Lys Val Leu Ala Gly Lys Ser Lys Ile Glu  
     770                      775                      780  
 Asp Tyr Phe Pro Glu Phe Ala Arg Tyr Thr Thr Pro Glu Asp Ala Thr  
     785                      790                      795                      800  
 Pro Glu Pro Gly Glu Asp Pro Arg Val Thr Arg Ala Lys Tyr Phe Ile  
                     805                      810                      815  
 Arg Asp Glu Phe Leu Arg Ile Ser Thr Ala Ser Gly Asp Gly Arg His  
                     820                      825                      830  
 Tyr Cys Tyr Pro His Phe Thr Cys Ala Val Asp Thr Glu Asn Ile Arg  
                     835                      840                      845  
 Arg Val Phe Asn Asp Cys Arg Asp Ile Ile Gln Arg Met His Leu Arg  
     850                      855                      860  
 Gln Tyr Glu Leu Leu  
     865

<210> 101  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 101  
 tctagaatga cgtccacctg caccaacagc

30

<210> 102  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 102  
 gatatcgtag gaaaagtagc agaatcgtag gaag

34

<210> 103  
 <211> 2781  
 <212> DNA  
 <213> Homo Sapiens and Rat

<400> 103  
 atgacgtcca cctgcaccaa cagcacgcgc gagagtaaca gcagccacac gtgcatgccc 60  
 ctctccaaaa tgcccatcag cctggcccac ggcacatcc gctcaaccgt gctggttatt 120  
 ttctctgccc cctctttcgt cggcaacata gtgctggcgc tagtggttgca gcgcaagccg 180  
 cagctgctgc aggtgaccaa ccgtttttatc tttaacctcc tcgtcaccga cctgctgcag 240

atttcgctcg	tgccccctg	ggtggtggcc	acctctgtgc	ctctcttctg	gcccccaac	300
agccacttct	gcacggccct	ggttagcctc	accacactgt	tcgccttcgc	cagcgtaac	360
accattgtcg	tgggtgcagt	ggatcgctac	ttgtccatca	tccacctct	ctcctaccg	420
tccaagatga	cccagcgccg	cggttacctg	ctcctctatg	gcacctggat	tgtggccatc	480
ctgcagagca	ctcctccact	ctacggctgg	ggccaggctg	cctttgatga	gcgcaatgct	540
ctctgctcca	tgatctgggg	ggccagcccc	agctacacta	ttctcagcgt	ggtgtccttc	600
atcgtcattc	cactgattgt	catgattgcc	tgctactccg	tgggtgttctg	tgagccccg	660
aggcagcatg	ctctgctgta	caatgtcaag	agacacagct	tggaagtgcg	agtcaaggac	720
tgtgtggaga	atgaggatga	agagggagca	gagaagaagg	aggagttcca	ggatgagagt	780
gagtttcgcc	gccagcatga	aggtgaggtc	aaggccaagg	agggcagaat	ggaagccaag	840
gacggcagcc	tgaaggccaa	ggaaggaagc	acggggacca	gtgagagtag	tgtagaggcc	900
aggggcagcg	aggaggtcag	agagagcagc	acgggtggcca	gcgacggcag	catggagggt	960
aaggaaggca	gcaccaaagt	tgaggagaac	agcatgaagg	cagacaaggg	tcgcacagag	1020
gtcaaccagt	gcagcattga	cttgggtgaa	gatgacatgg	agtttggtga	agacgacatc	1080
aatttcagtg	aggatgacgt	cgaggcagtg	aacatcccgg	agagcctccc	accagtcgt	1140
cgtaacagca	acagcaacct	tcctctgccc	aggtgctacc	agtgcaaagc	tgctaaagtg	1200
atcttcatca	tcattttctc	ctatgtgcta	tccttggggc	cctactgctt	tttagcagtc	1260
ctggccgtgt	gggtggatgt	cgaaaccag	gtaccccgat	gggtgatcac	cataatcatc	1320
tggcttttct	tcctgcagtg	ctgcatccac	ccctatgtct	atggctacat	gcacaagacc	1380
attaagaagg	aatccagga	catgctgaag	aagttcttct	gcaaggaaaa	gccccgaaa	1440
gaagatagcc	accgagacct	gcccggaaac	gagggtggga	ctgaaggcaa	gattgtccct	1500
tcctacgatt	ctgctacttt	tcctgcgata	tctgcagaat	tccaccacac	tgactagtgt	1560
gatccgagct	cggtagcaag	cttgggtgct	aggtcgatgg	gctgcctcgg	caacagtaag	1620
accgaggacc	agcgcaacga	ggagaaggcg	cagcgcgagg	ccaacaaaaa	gatcgagaag	1680
cagctgcaga	aggacaagca	ggtctaccgg	gccacgcacc	gcctgctgct	gctgggtgct	1740
ggagagtctg	gcaaaagcac	cattgtgaag	cagatgagga	tcctacatgt	taatgggttt	1800
aacggagagg	gcggcgaaga	ggacccgcag	gctgcaagga	gcaacagcga	tggtgagaag	1860
gccaccaaag	tgcaggacat	caaaaacaac	ctgaaggagg	ccattgaaac	cattgtggcc	1920
gccatgagca	acctggtgcc	ccccgtggag	ctggccaacc	ctgagaacca	gttcagagtgt	1980
gactacattc	tgagcgtgat	gaacgtgcc	aactttgact	tcccacctga	attctatgag	2040
catgccaaag	ctctgtggga	ggatgaggga	gttcgtgcct	gctacgagcg	ctccaacgag	2100
taccagctga	tcgactgtgc	ccagtacttc	ctggacaaga	ttgatgtgat	caagcaggcc	2160
gactacgtgc	caagtgaacca	ggacctgctt	cgctgccgcg	tcctgacctc	tggaatcttt	2220

gagaccaagt tccaggtgga caaagtcaac ttccacatgt tcgatgtggg cggccagcgc 2280  
 gatgaacgcc gcaagtggat ccagtgttc aatgatgtga ctgccatcat ctctgtggtg 2340  
 gccagcagca gctacaacat ggtcatccgg gaggacaacc agaccaaccg tctgcaggag 2400  
 gctctgaacc tcttcaagag catctggaac aacagatggc tgcgtaccat ctctgtgatc 2460  
 ctcttctca acaagcaaga tctgcttgct gagaaggtcc tcgctgggaa atcgaagatt 2520  
 gaggactact ttccagagtt cgctcgctac accactcctg aggatgagac tcccagagccc 2580  
 ggagaggacc cacgcgtgac ccgggccaag tacttcatcc gggatgagtt tctgagaatc 2640  
 agcactgcta gtggagatgg acgtcactac tgctacctc actttacctg cgccgtggac 2700  
 actgagaaca tccgccgtgt cttcaacgac tgccgtgaca tcatccagcg catgcatctt 2760  
 cgccaatacg agctgctcta a 2781

<210> 104  
 <211> 926  
 <212> PRT  
 <213> Homo sapiens and Rat

<400> 104

Met	Thr	Ser	Thr	Cys	Thr	Asn	Ser	Thr	Arg	Glu	Ser	Asn	Ser	Ser	His
1				5					10					15	
Thr	Cys	Met	Pro	Leu	Ser	Lys	Met	Pro	Ile	Ser	Leu	Ala	His	Gly	Ile
			20					25					30		
Ile	Arg	Ser	Thr	Val	Leu	Val	Ile	Phe	Leu	Ala	Ala	Ser	Phe	Val	Gly
	35						40					45			
Asn	Ile	Val	Leu	Ala	Leu	Val	Leu	Gln	Arg	Lys	Pro	Gln	Leu	Leu	Gln
	50					55					60				
Val	Thr	Asn	Arg	Phe	Ile	Phe	Asn	Leu	Leu	Val	Thr	Asp	Leu	Leu	Gln
	65				70					75					80
Ile	Ser	Leu	Val	Ala	Pro	Trp	Val	Val	Ala	Thr	Ser	Val	Pro	Leu	Phe
			85						90					95	
Trp	Pro	Leu	Asn	Ser	His	Phe	Cys	Thr	Ala	Leu	Val	Ser	Leu	Thr	His
			100					105					110		
Leu	Phe	Ala	Phe	Ala	Ser	Val	Asn	Thr	Ile	Val	Val	Val	Ser	Val	Asp
	115						120					125			
Arg	Tyr	Leu	Ser	Ile	Ile	His	Pro	Leu	Ser	Tyr	Pro	Ser	Lys	Met	Thr
	130					135					140				
Gln	Arg	Arg	Gly	Tyr	Leu	Leu	Leu	Tyr	Gly	Thr	Trp	Ile	Val	Ala	Ile
	145				150					155					160
Leu	Gln	Ser	Thr	Pro	Pro	Leu	Tyr	Gly	Trp	Gly	Gln	Ala	Ala	Phe	Asp
				165					170					175	
Glu	Arg	Asn	Ala	Leu	Cys	Ser	Met	Ile	Trp	Gly	Ala	Ser	Pro	Ser	Tyr
			180					185					190		

Thr Ile Leu Ser Val Val Ser Phe Ile Val Ile Pro Leu Ile Val Met  
 195 200 205  
 Ile Ala Cys Tyr Ser Val Val Phe Cys Ala Ala Arg Arg Gln His Ala  
 210 215 220  
 Leu Leu Tyr Asn Val Lys Arg His Ser Leu Glu Val Arg Val Lys Asp  
 225 230 235 240  
 Cys Val Glu Asn Glu Asp Glu Glu Gly Ala Glu Lys Lys Glu Glu Phe  
 245 250 255  
 Gln Asp Glu Ser Glu Phe Arg Arg Gln His Glu Gly Glu Val Lys Ala  
 260 265 270  
 Lys Glu Gly Arg Met Glu Ala Lys Asp Gly Ser Leu Lys Ala Lys Glu  
 275 280 285  
 Gly Ser Thr Gly Thr Ser Glu Ser Ser Val Glu Ala Arg Gly Ser Glu  
 290 295 300  
 Glu Val Arg Glu Ser Ser Thr Val Ala Ser Asp Gly Ser Met Glu Gly  
 305 310 315 320  
 Lys Glu Gly Ser Thr Lys Val Glu Glu Asn Ser Met Lys Ala Asp Lys  
 325 330 335  
 Gly Arg Thr Glu Val Asn Gln Cys Ser Ile Asp Leu Gly Glu Asp Asp  
 340 345 350  
 Met Glu Phe Gly Glu Asp Asp Ile Asn Phe Ser Glu Asp Asp Val Glu  
 355 360 365  
 Ala Val Asn Ile Pro Glu Ser Leu Pro Pro Ser Arg Arg Asn Ser Asn  
 370 375 380  
 Ser Asn Pro Pro Leu Pro Arg Cys Tyr Gln Cys Lys Ala Ala Lys Val  
 385 390 395 400  
 Ile Phe Ile Ile Ile Phe Ser Tyr Val Leu Ser Leu Gly Pro Tyr Cys  
 405 410 415  
 Phe Leu Ala Val Leu Ala Val Trp Val Asp Val Glu Thr Gln Val Pro  
 420 425 430  
 Gln Trp Val Ile Thr Ile Ile Ile Trp Leu Phe Phe Leu Gln Cys Cys  
 435 440 445  
 Ile His Pro Tyr Val Tyr Gly Tyr Met His Lys Thr Ile Lys Lys Glu  
 450 455 460  
 Ile Gln Asp Met Leu Lys Lys Phe Phe Cys Lys Glu Lys Pro Pro Lys  
 465 470 475 480  
 Glu Asp Ser His Pro Asp Leu Pro Gly Thr Glu Gly Gly Thr Glu Gly  
 485 490 495  
 Lys Ile Val Pro Ser Tyr Asp Ser Ala Thr Phe Pro Ala Ile Ser Ala  
 500 505 510  
 Glu Phe His His Thr Gly Leu Val Asp Pro Ser Ser Val Pro Ser Leu  
 515 520 525  
 Gly Cys Arg Ser Met Gly Cys Leu Gly Asn Ser Lys Thr Glu Asp Gln  
 530 535 540

Arg Asn Glu Glu Lys Ala Gln Arg Glu Ala Asn Lys Lys Ile Glu Lys  
 545 550 555 560  
 Gln Leu Gln Lys Asp Lys Gln Val Tyr Arg Ala Thr His Arg Leu Leu  
 565 570 575  
 Leu Leu Gly Ala Gly Glu Ser Gly Lys Ser Thr Ile Val Lys Gln Met  
 580 585 590  
 Arg Ile Leu His Val Asn Gly Phe Asn Gly Glu Gly Gly Glu Glu Asp  
 595 600 605  
 Pro Gln Ala Ala Arg Ser Asn Ser Asp Gly Glu Lys Ala Thr Lys Val  
 610 615 620  
 Gln Asp Ile Lys Asn Asn Leu Lys Glu Ala Ile Glu Thr Ile Val Ala  
 625 630 635 640  
 Ala Met Ser Asn Leu Val Pro Pro Val Glu Leu Ala Asn Pro Glu Asn  
 645 650 655  
 Gln Phe Arg Val Asp Tyr Ile Leu Ser Val Met Asn Val Pro Asn Phe  
 660 665 670  
 Asp Phe Pro Pro Glu Phe Tyr Glu His Ala Lys Ala Leu Trp Glu Asp  
 675 680 685  
 Glu Gly Val Arg Ala Cys Tyr Glu Arg Ser Asn Glu Tyr Gln Leu Ile  
 690 695 700  
 Asp Cys Ala Gln Tyr Phe Leu Asp Lys Ile Asp Val Ile Lys Gln Ala  
 705 710 715 720  
 Asp Tyr Val Pro Ser Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr  
 725 730 735  
 Ser Gly Ile Phe Glu Thr Lys Phe Gln Val Asp Lys Val Asn Phe His  
 740 745 750  
 Met Phe Asp Val Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln  
 755 760 765  
 Cys Phe Asn Asp Val Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser  
 770 775 780  
 Tyr Asn Met Val Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu  
 785 790 795 800  
 Ala Leu Asn Leu Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr  
 805 810 815  
 Ile Ser Val Ile Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys  
 820 825 830  
 Val Leu Ala Gly Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala  
 835 840 845  
 Arg Tyr Thr Thr Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro  
 850 855 860  
 Arg Val Thr Arg Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile  
 865 870 875 880  
 Ser Thr Ala Ser Gly Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr



885 890 895  
 Cys Ala Val Asp Thr Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg  
           900                  905                  910

Asp Ile Ile Gln Arg Met His Leu Arg Gln Tyr Glu Leu Leu  
           915                  920                  925

<210> 105  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 105  
 catgtatgcc agcgtcctgc tcc

23

<210> 106  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 106  
 gctatgcctg aagccagtct tgtg

24

<210> 107  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 107  
 gcacctgctc ctgagcacct tctcc

25

<210> 108  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_feature  
 <223> Novel Sequence

<400> 108  
 cacagcgctg cagccctgca gctggc

26

<210> 109

<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 109  
ccagtgatga ctctgtccag cctg

24

<210> 110  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 110  
cagacacttg gcagggacga ggtg

24

<210> 111  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 111  
cttgtggtct actgcagcat gttccg

26

<210> 112  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 112  
catatccctc cgagtgtcca gcggc

25

<210> 113  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc feature  
<223> Novel Sequence

<400> 113  
atggatcctt atcatggctt cctc

24

<210> 114  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 114  
caagaacagg tctcatctaa gagctcc

27

<210> 115  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 115  
ctctgatgcc atctgctgga ttcctg

26

<210> 116  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 116  
gtagtccact gaaagtccag tgatcc

26

<210> 117  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 117  
tggtggcgat ggccaacagc gctc

24

<210> 118  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 118  
gttgcgccctt agcgacagat gacc

24

<210> 119  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 119  
tcaacctgta tagcagcatc ctc

23

<210> 120  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 120  
aaggagtagc agaattggta gcc

23

<210> 121  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 121  
gacacctgtc agcggtcgtg tgtg

24

<210> 122  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 122  
ctgatggaag tagaggctgt ccatctc

27

<210> 123  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 123  
gcgctgagcg cagaccagtg gctg

24

<210> 124  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 124  
cacggtgacg aagggcacga gctc

24

<210> 125  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 125  
agccatccct gccaggaagc atgg

24

<210> 126  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 126  
ccaggtaggt gtgcagcaca atggc

25

<210> 127  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 127  
ctgttcaaca gggctggttg gcaac

25

<210> 128  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 128  
atcatgtcta gactcatggt gatcc

25

<210> 129  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 129  
Thr Leu Glu Ser Ile Met  
1 5

<210> 130  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 130  
Glu Tyr Asn Leu Val  
1 5

<210> 131  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> misc\_feature  
<223> Novel Sequence

<400> 131  
Asp Cys Gly Leu Phe  
1 5

<210> 132

<211> 36  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 132

Gly Ala Thr Cys Ala Ala Gly Cys Thr Thr Cys Cys Ala Thr Gly Gly  
 1 5 10 15

Cys Gly Thr Gly Cys Thr Gly Cys Cys Thr Gly Ala Gly Cys Gly Ala  
 20 25 30

Gly Gly Ala Gly  
 35

<210> 133  
 <211> 53  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> misc feature  
 <223> Novel Sequence

<400> 133

Gly Ala Thr Cys Gly Gly Ala Thr Cys Cys Thr Thr Ala Gly Ala Ala  
 1 5 10 15

Cys Ala Gly Gly Cys Cys Gly Cys Ala Gly Thr Cys Cys Thr Thr Cys  
 20 25 30

Ala Gly Gly Thr Thr Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly Ala  
 35 40 45

Thr Gly Gly Thr Gly  
 50